Higher Education Expansion and the Rise of the Service Economy in Vietnam*

(Click here for the latest version)

Khoa Vu†

Abstract

Over the last two decades, developing countries have witnessed the rise of the service sector, raising concerns about pre-mature deindustrialization. We explore expansion of access to higher education as a potential driver behind this trend, using Vietnam as a case study. We leverage a national expansion of higher education in Vietnam, which established over 100 universities from 2006 to 2013. Collecting a dataset on the timing and location of university openings, we estimate that individual's exposure to a university opening increases their chances of completing college by over 57%, allowing workers to move from the agricultural sector into service sector. At the firm-level, exposed service firms experience a significant increase in productivity. We show that these findings are consistent with a model of endogenous technological change in response to an increase in skilled labor supply.

Keywords: higher education, local labor market, labor productivity, Vietnam

*The authors thank Paul Glewwe, Marc Bellemare, Jason Kerwin, Aaron Sojourner, Joe Ritter, Ian Coxhead, Duc H. Vu, Luis Baldomero-Quintana, Sunghun Lim, Nguyen Vuong, Dave Evans, Abhijeet Singh, Quoc-Anh Do, Esteban J. Quiñones, Mike Andrews, Saurabh Singhal, the audience members in the ERIA workshop, Paul Glewwe's lab, the Bush School of Government and Public Service seminar, the Grand Valley State University Economics Seminar, and the Armenian Economic Association Seminar for their valuable comments. All errors are of the authors.

†Khoa Vu (corresponding author). Email: vuxxx121@umn.edu. Department of Applied Economics, University of Minnesota.

1 Introduction

Many developing countries have recently experienced a substantial increase in the employment of the service sector with important implications for the economic development process. On the one hand, these countries are skipping entirely the industrialization phase, a phenomenon known as pre-mature deindustrialization, which can potentially be detrimental to long-term growth (Rodrik, 2016; McMillan et al., 2017). On the other hand, the rise of the service sector may signal a productivity growth in the service sector, which can be an important driver of long-term growth (Buera et al., 2022; Fan et al., 2023; Chen et al., 2023). Understanding what drives this servicification process is, thus, an important question for both researchers and policymakers.[1]

Coinciding with the rise of the service sector is a rapid expansion of access to higher education in the developing countries, as the enrollment rates in the low-income countries have more than doubled for the last twenty years (The World Bank, 2022). One may ask whether the expansion of higher education is partly responsible for the 'servicification' process?[2] Since the service sector is skill-intensive, an increase in the supply of college-educated workers leads to the labor reallocation from agriculture to service, which, in turn, drives the increase in productivity of firms in the service sector as firms adopt better technologies. Yet it is not hard to imagine that the causality is the other way around: the expansion of the service sector leads to higher income, allowing more individuals to afford attending university.

To answer this question, we leverage a national expansion of higher education in Vietnam to study the link between access to higher education and the rise of the service sector in

[1] Rodrik (2016) provides suggestive evidence that globalization and trade are the main factor behind this trend. Fan et al. (2023) and Chen et al. (2023) use a structural model to show that the rises of the service sector in India and China are driven by the increase of productivity in this sector, though it remains unclear what drives this productivity increase.

[2] Note that this is different from the rise of the service economy in developed countries, which is often characterized by increases in both skilled labor and skill premium (Buera and Kaboski, 2012; Buera et al., 2022).

terms of employment and productivity at the worker- and firm-levels. Vietnam is in a unique position to help us understand the causal link between these two trends. In 2006, the government issued Decree 121/2007/QD-TTg, which led to the establishment of over 100 new universities across the country between 2006 and 2013; the proportion of young adults with a college education almost tripled during the same period (see Figure 2). This happened against the back drop of rapid labor reallocation from agriculture to manufacturing that started in the late 1980s (McCaig and Pavcnik, 2013), whereas employment in the service sector only started to rise in 2010, four years after the expansion started.

We collect a new dataset on the timing and locations of all university openings in Vietnam,[3] leveraging the staggered timing of new university establishments to understand how workers and firms are affected. In the first part of the paper, we examine whether expanding access to higher education leads to structural transformation at the worker-level, i.e., whether workers exposed to better access to higher education are more likely to work in sectors with higher productivity, e.g., manufacturing and service. We estimate a difference-in-differences model comparing across birth cohorts and provinces and find that being exposed to a new university at college-going age decreases the likelihood of working in the agricultural sector and increases the likelihood of working in the service sector. Exposed workers are found to have higher wages.

Our results are robust when combining the two-way fixed effects model with propensity score matching and to accounting for differential linear trends across the treatment and control groups. We further show that migration does not explain our results, nor does the expansion trigger any migration response across provinces. The effects across most outcomes are concentrated among female workers, especially for the average monthly wage. Using a change-in-changes (CiC) model, we document substantial treatment effect heterogeneity on

[3]All public and private university openings require official approvals from the government, which are publicly available and contain information about the location and timing of university openings. We hand-entered these information for all universities established in Vietnam since 1975.

wage; the treatment effects are particularly high at the bottom of the wage distribution, where the share of college-educated workers is very low.

In the second part of the paper, we explore the effects on firms through changes in the local labor market. Using a similar staggered timing difference-in-differences model, I first show that the expansion increases the relative supply of college-educated workers and lowers the college wage premium. At the firm-level, I find that the expansion leads to higher productivity in the service sector in the long run. This result is consistent when being measured by total factor productivity or labor productivity (revenue per worker). We show that these results are consistent with a canonical model of skill differentials with different age cohorts (Card and Lemieux, 2001) and endogenous technological adoption (Acemoglu, 1998; Blundell et al., 2022; Carneiro et al., 2023).

These results indicate that expanding access to higher education can indeed lead to a rise of employment and productivity in the service sector in Vietnam, making our study related to three broad bodies of literature. First, a large macroeconomic literature documents the important role of human capital in the structural transformation process (e.g., Schoellman, 2012; Porzio et al., 2022; Hendricks and Schoellman, 2023; Martellini et al., 2022). A separate literature focuses on the rise of the service sector in developing countries (Rodrik, 2016; McMillan et al., 2017; Fan et al., 2023; Chen et al., 2023), which is driven either by globalization or the increase of productivity in the service sector. Our results point to the rise of higher education as an important factor of this type of structural transformation.

Second, our study contributes to a large literature on higher education as a driver of technological change and productivity growth,[4] which has been mainly studied in developed countries. An increase in college-educated workers can affect economic growth through the R&D market or productivity spillover.[5] A relatively recent literature established the impor-

[4]See Valero (2021) for an extensive review.

[5]See, e.g., Acemoglu (1998); Moretti (2004a,b); Vandenbussche et al. (2006).

tance of colleges in developed countries such as the United States and Norway, as they drive local educational attainment (Russell et al., 2022), innovation (Hausman, 2020; Andrews, 2020), technological adoption (Carneiro et al., 2023; Blundell et al., 2022), consumption (Liu and Yang, 2021), and agglomeration (Liu, 2015).

Third, our study contributes to a long literature on the economic impacts of providing access to education in developing countries, which mainly focuses on primary school education (e.g., Duflo, 2001; Akresh et al., 2018; Khanna, 2023)[6].[7] A number of studies exploring the effects of a higher education expansion in China in 1999 on labor market outcomes,[8] firm productivity (Che and Zhang, 2018), and technological adoption (Feng and Xia, 2022). Unlike the expansion in Vietnam, which established new universities, the expansion in China raised the college admission quotas (Li et al., 2017), which also raised college enrollment and attainment. The key difference between our study and these studies is that we can exploit the variation in locations and timing of new universities, while the other studies rely on enrollment rates or indirect measures of exposure such as rural/urban status.

The rest of the paper is organized as follows. In Section 2 and 3, we briefly describe about the national expansion of higher education and the data sources that we use. In Section 4, we discuss the empirical strategy to evaluate the effects of individual exposure to the expansion and the results from our estimations. In Section 5, we present a capital-skill complementarity model to understand how the expansion may affect firms through the labor market. We then discuss how we evaluate the effects on these outcomes and the results from our estimations. We discuss the implications and shortcomings of our study in Section 7.

[6]This literature includes a small but growing literature on education in Vietnam. Dang and Glewwe (2018) and Dang et al. (2021) explore Vietnam's exceptional performance in basic education relative to its past and other countries. Phan and Coxhead (2013) examine the economic forces behind changes in returns to schooling. Coxhead and Shrestha (2017) study how foreign direct investment affects schooling decisions. Several studies examine the overall changes in returns to schooling in Vietnam (e.g., Patrinos et al., 2018; Doan et al., 2018; McGuinness et al., 2021)

[7]A related literature studies whether the value of higher education in developing countries is due to human capital or signaling (Arteaga, 2018; Barrera-Osorio and Bayona-Rodríguez, 2019).

[8]See, e.g., Li et al. (2014, 2017,?); Xing et al. (2018); Huang et al. (2022)

2 Background of the Higher Education Expansion in Vietnam

The Decree 121/2007, signed by the Prime Minister Nguyen Tan Dung, laid out the overall plan of the government to expand the network of higher education institutions across the country during the 2006-2020 period. As a result of this plan, the number of universities in Vietnam went from 150 in 2005 to over 250 in 2014 (see Figure 2). In 2013, the government announced that it had already reached the number of universities planned for this period, so no new universities were established after that (with a few exceptions).[9]

Universities are also distributed unevenly across 63 provinces of Vietnam, as shown in Figure 3. Ho Chi Minh City, Hanoi, and Hue were the three "centers" before the expansion, with more than 6 universities in each city/province. After the expansion, other provinces also saw an increase in the number of universities, such as Thai Nguyen and Nghe An in the North and Binh Duong and Dong Nai in the South. Many provinces had the first university open during this period.

To understand the impacts of these new universities, we collect data on all universities (including existing universities) from official documents of the government. Both public and private universities were established during this expansion. While public universities were typically "upgraded" from two-year colleges, private universities were usually newly established rather than upgraded. Public universities are established by officials at the province level. However, both types of universities have to justify to the government that they have enough staff, faculty, infrastructure, and land to operate a university.

New universities tend to be opened in provinces with a large proportion of workers in skill-intensive sectors. We present in Table 1 provincial characteristics (before treatment) for three

[9]See Parajuli et al. (2020) and Vu and Nguyen (2018) for overviews of this policy.

groups of provinces: those that never had any university during our study period (our control group), those that opened a university for the first time (our treatment group), and provinces that already have a university, i.e., the "already-treated" group. The treatment group has a smaller share of agricultural workers and larger shares of workers in the manufacturing and service sectors. It is also unsurprising to see that the treatment group has a lower poverty rate than the control group. Besides these characteristics, the treatment and the control groups are relative similar in terms of other economic indicators such as employment rate, self-employment rate, and firm-level performances such as total factor productivity, labor productivity, and capital-labor ratio.

It is worth pointing out that the already-treated provinces are substantially different from the control group. The adult population tends to have better education, which is expected given that there are already existing universities. These provinces also have lower percentage of self-employment, higher share of employment, higher shares of manufacturing and service workers, and better economic conditions.

These results reflect the nonrandom nature of where new universities are opened. Provinces that open new universities likely respond to growing demands for higher education and a growing workforce in industries that require more education. However, the overall economic conditions are more similar to those in the control group than the already-treated group. These results strongly suggest that simply comparing across provinces may yield biased estimates because of selection into establishing new universities.

3 Data

Our study draws from several data sources. Data on individual and market-level labor market outcomes are based on the Labor Force Survey (LFS). We use the individual-level data for 2010-2018 to examine the individual benefits of being exposed to the higher education expansion. For the market-level analysis, we aggregate this data at the district-by-year level. It is also important to note that these data have information on districts only for 2011 and 2015-2019. This limits our ability to estimate event-study specifications at the district level.

For the individual-level analysis, we focus on college education, monthly wage, and employment as the main outcomes of interest. The LFS asks respondents for their highest educational attainment, which we use to construct a binary variable for whether an individual has obtained a 4-year university degree or more. Employment is measured by whether respondents are employed, as opposed to being self-employed or not working. We also use whether individuals earn a wage as an alternative measure for employment. Monthly wages only include wage compensation. but not other benefits such as bonuses.[10] We have this information for most workers, including those who are self-employed.

For the market-level analysis in the second part of the paper, we aggregate data up to the district-by-year level. Our first outcome variable is the share of working-age adults who have completed college education. Guided by a capital-skill complementarity model that we discuss in Section 5.1, we examine the effects of the expansion on the relative supply of college-educated workers, $\ln\left(\frac{H}{L}\right)$. We measure H and L as the number of college-educated and non-college adults with non-zero monthly wage. The second main outcome of interest is the college wage premium, also known as the relative wage of college-educated workers, measured as the log ratio of college-educated monthly wage to non-college monthly wage, i.e., $\ln\left(\frac{w_H}{w_L}\right)$. All observations with zero wage are treated as missing and dropped before aggregation. For both types of analyses, we restrict our analyses to those who are between age 22 and 54 since they are most likely out of school and active in the labor market. Given that Hanoi and Ho Chi Minh city are both the centers of economic growth and universities of Vietnam, we also exclude them from the analyses.

Lastly, we study firms' response to the labor supply shock along three dimensions: total

[10]This is because questions about bonus are not consistent across years.

factor productivity (TFP), labor productivity, and capital intensity using firm-level data from the Vietnam Enterprise Census (VEC) for 2006-2018. This data contains detailed accounting information on firms' annual operations, such as short- and long-term assets, total labor, total revenue, and industries. We measure labor productivity by value added per labor, where value-added is defined as profit plus wage (Newman et al., 2015). To measure the TFP of each firm, we first estimate the relevant production function for all 2-digit industries using Ackerberg et al. (2015)'s approach then use the estimated parameters to obtain the total factor productivity (see Appendix A for a detailed discussion about the estimation process and results). We measure capital intensity as the ratio of capital to revenue.

4 Impacts on Individual-Level Outcomes

In this section, we explore the effects of the expansion on individual-level outcomes. The expansion increases accessibility to higher education for younger cohorts; specifically, those who were at the college-going age when there were more new universities because of the reform. In contrast, it does not affect older cohorts who would generally be too old to benefit from the increase in number of universities. As the younger cohorts have better access to higher education, they may be more likely to obtain a university degree, and hence, may do better than older cohorts in terms of employment and wages. We discuss this difference-in-differences design and estimation in Section 4.1 and the results in Section 4.2. We further explore the effects of being exposed to the expansion on the type of occupation and the industry of work.

In Section **??**, we discuss why the cohort-based DiD estimates do not identify the partial equilibrium returns to higher education given the likely general equilibrium (GE) forces in our context, and how one can estimate such returns if knowing the GE effects of the expansion.

4.1 Empirical Strategy

We take advantage of the variation in the opening dates and locations of these universities to identify the effects of the expansion on workers and firms. At the worker level, we compare birth cohorts that were 25 years old or younger when the expansion took place and, thus, would benefit from having access to new universities (i.e., the *exposed* cohorts), and birth cohorts that were older than 25 years old and, thus, would be too old to benefit from the expansion (i.e., *unexposed* cohorts). We then compare the cohort differences across provinces that established universities for the first time (treatment provinces) and those that never established a university before (control provinces). There are two reasons for choosing 25 as the cutoff age. Those who are slightly older than 18 years old may still be eligible for college (e.g., those who repeated a class in earlier ages). Another reason is that individuals may choose to attend a college (2-year or 3-year degree) before transferring to a university.

Given that the timing of university opening varies across provinces, let g denote the year that a given province has its first university ever, t the survey year, and c the birth year of a given cohort. Let G_g denote the group of provinces with the same treatment year. For each group of provinces with the same treatment year G_g in the same survey year, we can compare them to the never-treated provinces as the control group across the exposed and unexposed birth cohorts.

That is, we can estimate the following two-way fixed effects (TWFE) model in a sub-dataset with only provinces in G_g (treatment group with treatment year g) and the never-treated provinces (control group) in the same survey year t:

$$y_{i,p,c}^{G,t} = \delta^{G,t}(T_p^{G,t} \times Exposed_c^{G,t}) + \gamma_p + \eta_c + \epsilon_{i,p,c}$$

where $y_{i,p,c}$ denotes the outcome of individual i in province p of cohort c; T_p indicates whether province p had the first university during the expansion; $Exposed_c$ indicates whether cohort

c was 21 years old or younger when the first university was opened in year g. Province and cohort fixed effects are γ_p and η_c. Since we are comparing each treatment group to a never-treated group in a given year, $\delta^{G,t}$ captures the treatment effect of a specific group G in a specific year t. One can then aggregate $\delta^{G,t}$ across all G and t to obtain the average effect for different groups of provinces and for different years.

Alternatively, we can combine all sub datasets and estimate the following model:

$$y_{i,p,c,s} = \delta(T_{p,s} \times Exposed_{c,p,s}) + \gamma_{p,s} + \eta_{c,s} + \epsilon_{i,p,c,s} \tag{1}$$

where s denotes the subdataset for each G and t, while $\gamma_{p,s}$ and $\eta_{c,s}$ control for province-by-subdataset and cohort-by-subdataset fixed effects. Thus, δ captures the weighted average of all $\delta^{G,t}$. To ensure that we are capturing the effect of having access to higher education via the expansion, we limit our analysis to survey years that are at least 4 years after the treatment year g, allowing enough time for the exposed cohorts to complete their 4-year degree before entering the labor market.

This approach, proposed by Cengiz et al. (2019), allows us to compare each treatment group G with a clean control group in each given year, thus avoiding the problem of negative weights in a standard TWFE model (De Chaisemartin and d'Haultfoeuille, 2020; Goodman-Bacon, 2021). There are, of course, other different estimators such as Callaway and Sant'Anna (2021), Borusyak et al. (2021), and others. Yet this estimator is particularly suitable given that we only have repeated cross-sectional data and need to handle treatment effect heterogeneity across more than two dimensions.

The DiD model imposes a conditional parallel trends assumption, i.e., without the expansion, the outcome variables would have evolved similarly for the treatment and control provinces, conditional on the fixed effects and control variables. For the rest of this section, we discuss when this assumption might be violated and other concerns about estimation and inference strategy, as well as how we address these concerns.

4.1.1 Threats to the Parallel Trends Assumption

Three important threats to the parallel trends assumption are considered here. First, the treatment group and the control group might have followed pre-expansion differential trends, which may bias our results. Second, given that provinces that opened the first university and those that did not may be systematically different, and the differences might have differential effects across birth cohorts. Lastly, our results might be confounded by inter-provincial migration. Either one of these threats can lead to the parallel trends assumption being violated.

To address the concern about pre-expansion trends, we extend our stacked regression model to estimate an event study specification that allows us to visualize the parallel pre-trends (or lack thereof). Following Bilinski and Hatfield (2018)'s proposed test for parallel pre-trends, we assume that there is no differential trends between the treatment and the control groups in the true model. Thus, if we allow the two groups to trend differently by adding an interaction term $T_{p,s} \times \text{Age Exposed}_c$ in the TWFE model, then the estimated treatment effects of such a specification and the base specification should be similar.

To address a related concern about non-random assignment of the treatment status, we first use propensity score matching (PSM) on the pre-expansion observables to select a more comparable control group before estimating the TWFE model. It is, however, important to note that such a matching approach is valid only when the pre-treatment covariates are strongly serially correlated. Matching on covariates that are moderately serially correlated in a difference-in-differences model can introduce bias due to regression to the mean (Daw and Hatfield, 2018). Therefore, we use propensity score matching only based on urban and poverty rate, which are both relatively more stable over time than other covariates.

In the Appendix, we also construct an alternative control group based on the not-yet-treated provinces. Given that the already-treated provinces and the not-yet-treated provinces only differ in the timing, we can assume that they are relatively more similar than the already-treated and the never-treated provinces. For every group G_g in year t, we construct a control group that is made up of provinces with new university for the first time 3 years ago or later. That is, the control group are provinces with new university founded in year g' where $t - g' \leq 3$.

A more severe threat to our identification strategy is inter-province migration. New universities increase the supply of college-educated workers, thus reducing their wage level at the market level. This may induce college-educated workers to migrate to other provinces. Alternatively, new universities may create more job opportunities for college-educated workers and attract even more of these workers. In other words, our results might be driven by the educational attainment of the out-migrants and in-migrants. We address this issue by running the same regression on province of origin instead of province of current residence. The results, therefore, should not be affected by changes in the composition of provinces of current residence.

For the continuous monthly wage variable, another useful check is a change-in-changes model (Athey and Imbens, 2006) that imposes a weaker assumption than a parallel trends assumption. Specifically, it allows the distribution of outcome to vary in terms of mean and variance in the absence of treatment. The identifying assumption of this model is that in the absence of treatment, the distribution of the unobservables can vary between the treatment and control groups, but not across time within group (Athey and Imbens, 2006; Imbens and Wooldridge, 2009). If this alternative model yields a similar result to that of the standard DiD model even though it does not rely on the parallel trends assumption, we can be more confident that our results are not driven by a violation of that assumption.

4.2 Results for Worker-Level Effects

In Table 2, we show the descriptive statistics of the relevant variables for this analysis. Specifically, we present the mean and standard deviation for the unexposed cohorts (i.e., exposed between age 26 to 36) and the exposed cohorts in the control and treatment provinces.

In Table 3, we present the results from estimating the difference-in-differences model leveraging the variation in exposure across birth cohorts, provinces, and survey year. In the first specification, we estimate a standard TWFE model with a stacked regression as discussed in Section 4.1. In specification 2, we first use propensity score matching (PSM) to select a more comparable control group based on pre-treatment observable characteristics, specifically share of workers in different sectors as well as average income, urban status, and poverty rate. In the third specification, we estimate the main TWFE model but allow the treatment and control provinces to have differential linear trends. In specification 3, we estimate a similar model as the first specification, but including an interaction term for treatment status and linear trends to allow the treatment and control groups to trend differently. In all specifications, we allow treatment effects to vary among those who were exposed at age 18 or younger and those who were exposed between age 18 and 25.

We address migration in three different ways. In specification 4, we estimate the same TWFE model on a non-migrant subsample. In specification 5, we first construct a dataset with the treatment status based on province of origin instead of province of current residence. We then estimate a similar TWFE stacked regression on this dataset. If migration is the driver of our result, these two specifications would likely produce results that are different from the main models. Lastly, we estimate the effects of the expansion on inter-provincial migration status and report them in Figure A2.

We find that the expansion raises the probability of completing college of exposed individuals by 3.6 to 4.5 percentage points, as indicated by the results in column (1). Accounting for pre-treatment observables via PSM or differential linear trends brings the estimate slightly down, but the result is robust. Notably, those who were exposed between age 19 to 25 also saw an increase of 1.5 to 2.2 percentage points in college completion rate. This is likely due to potential spillover effect of college expansion as slightly older cohorts may also try to attend college as it becomes available in their provinces. The adjustments for differences in observables and differential pre-trends do not alter the result significantly, implying that our result is not driven by pre-treatment trends. Figure 5(a) confirms this implication. Those who were exposed to the expansion between age 36 and 27 saw very little effects on college completion, while those who were exposed in latter ages saw larger effects.

In column (2), we find that being exposed to the expansion between age 14 to 18 years old raises the probability of employment by 13.1 percentage points in the base specification. Adjusting for pre-treatment observables via PSM and differential trends lower the estimates to 12.6 percentage points and 7.4 percentage points, respectively. In other words, it can be concluded that the expansion raises employment for those who were exposed. The event study in Figure 5(b) yields a very similar conclusion.

Given the positive impact on employment, it is also unsurprising to find that exposure to the expansion also raises the monthly wage by 4 to 8.6 percentage points in column (3). The effects are significant in the main specification and in the one with PSM, but not in the third specification where we allow the treatment and control groups to have differential linear trends. However, Figure 5(c) does not indicate any significant pre-trends.

In columns (4) to (6), we examine whether the expansion has any effect on the sector of work. We find that exposed individuals are less likely to work in the agricultural sector, and more likely to work in the manufacturing and service sectors. It is, however, important to point out that the effect on manufacturing employment is reduced to almost zero when controlling for differential pre-trends, while the effects on the other sectors are robust to different specifications. The event study results in Figure 5(d), 5(e), and 5(f) yield very similar conclusions.

Taken together, the results indicate that exposure to the expansion increases the probability of completing college by more than 4 percentage points among those who were exposed at age 12 to 18; this is equivalent to an increase of over 57%. It also increases the chance of being employed of individuals, especially in the service sector. Subsequently, we also find that exposure to the expansion raises monthly wage by 4 to almost 9 percent.

Migration appears to play a very minor role in the effects of exposure. Specification 4 using a non-migrant sample generally yields similar results to the main specification, suggesting that the main findings are driven by individuals born in their province of current residence. Specification 5 using a sample with treatment status based on province of origin also confirms this conclusion, as the magnitudes of the estimations are similar to our main findings. In Figure A2, we estimate the effect of being exposed on individuals' migration status on two samples. The sample using treatment status based on province of current residence gives us the effect on in-migration, while the sample using treatment status based on province of origin gives us the effect on out-migration. The results on both types suggests that exposure to the expansion increases both migration flowing into and out of treatment provinces, but the treatment effects are small and almost always insignificant.

Figure A3 presents the distributional effects on log monthly wage using the change-in-changes (CiC) model (Athey and Imbens, 2006).[11] All treatment effects are positive and statistically significant; the treatment effects in most quantiles are also in line with the treatment effect estimated from the TWFE model above. Interesting, these results indicate that those at the bottom of the wage distribution see the largest increase in wage relative to those at the top.

These CiC results suggest that our wage estimates are partly driven by general equilibrium

[11] We use the Stata command *cic* (Melly and Santangelo, 2015) to estimate this CiC model.

effects. To see this, we also report the share of individuals completing college by wage quantile in the control provinces. The two lowest quantiles of the wage distribution have college completion rate less than 5%, while the treatment effects are higher than the rest of the distribution. In contrast, the two highest quantiles have up to 35% college completion rate, but the treatment effects are smaller. These results suggest that the non-college workers likely see a (larger) increase in wage due to the expansion.

4.3 Effects on Occupation and Industry of Work

An important implication of our findings is that being exposed to the expansion allows workers to move out of the agricultural sector and into more productive sectors such as service and manufacturing. This is a classic story of structural transformation of economic development: as workers become more educated, they move from less productive to more productive sectors, thus contributing to productivity growth (Porzio et al., 2022).

In this section, we take a closer look at the specific industries and occupations in which exposed individuals choose to work using the same cohort-based DiD model in Equation 3. The industry effects are reported in Figure A5. First, exposure to the expansion reduces the likelihood of working in agriculture, fishing, and forestry, as well as construction. Second, it increases the likelihood of working in education and health as well as politics. More surprisingly, exposure has a positive and significant effect on working in IT, finance, and science, as well as entertainment and other services, but the effect sizes are relatively small.

The occupation effects are reported in Figure A6. We observe that the expansion significantly reduces the likelihood of being an elementary worker (e.g., domestic helpers or manual labor). We also find positive effects on clerks (e.g., secretaries and bank tellers), technicians, scientists, experts, and managerial positions. The effect on being a scientist or expert is the most profound. These results strongly suggest that having access to higher education is valuable because it allows workers to obtain occupations and industries that typically require such a level of educational attainment.[12]

4.3.1 Effects by Gender

The CiC estimationd reveals interesting heterogeneity in wage effects across the wage distribution. One potential dimension in which this heterogeneity manifests itself is gender. Figure A7 provides the event study results estimated by gender. The effects on college education of females are relatively large, while those of males are smaller. As a result, the positive effects on labor market outcomes such as employment and wage are also larger among female than males. The wage effects among males are almost zero. These differences are also consistent with previous studies (Elsayed and Shirshikova, 2023).

There are three potential explanations for such a large gap in treatment effects by gender. First, our results indicate that the expansion has a much larger effect on service employment than other sectors. Since service employment is also dominated by women, we can think of the expansion as an intervention to lower the cost of obtaining a degree to get a career in the service sector, so this intervention is particularly more relevant for female than for male.

An alternative explanation is that parents are less willing to let daughters to travel far for college education compared to sons. Opening a new university in a province thus allows girls in that area to access to higher education without traveling. Sánchez and Singh (2018) also offer another alternative explanation: in Vietnam, aspiration for higher education of girls is higher than that of boys. Aspiration is also an important predictor of enrollment for girls but not for boys. Therefore, improving access unlocks the possibility that girls can enroll in higher education in Vietnam.

[12] In Figure A4, we show that exposed individuals are more likely to have formal employment, have higher-paid occupation, and work in more skill-intensive industries.

5 Effects on Firm Productivity

We now turn to examine the effects of the expansion on firm-level productivity. To guide our empirical exploration, we first consider a theoretical model of endogenous technological adoption by Acemoglu (2007) and Carneiro et al. (2023). In this model, the expansion increases the relative supply of college-educated workers and, thus, lowers the college wage premium. The skill-biased technology becomes relatively more profitable as the supply of skilled labor rises and firms adopt them once the supply crosses a certain threshold. We then extend the stacked regression approach to study the GE effects of the expansion on province-level labor market as well as firm-level productivity and capital intensity. Lastly, we revisit the partial equilibrium returns to higher education.

5.1 A Model of Endogenous Technological Adoption

We present a simple framework on the relationship between skill premium and relative supply of skilled workers (Card and Lemieux, 2001) in the presence of endogenous technological change (Acemoglu, 1998; Moretti, 2004a; Blundell et al., 2022; Carneiro et al., 2023; Khanna, 2023). First, let us denote two cohorts, $j = \{y, o\}$ y and two levels of education, $e = \{c, n\}$ as before. Suppose the following constant elasticity of substitution (CES) production function:

$$Y_t = [\alpha.(A_c.L_c)^{\frac{\sigma_E - 1}{\sigma_E}} + (1-\alpha).(A_n.L_n)^{\frac{\sigma_E - 1}{\sigma_E}}]^{\frac{\sigma_E}{\sigma_E - 1}}$$

where $L_c = (\phi_{cy} l_{cy}^{\frac{\sigma_A - 1}{\sigma_A}} + \phi_{co} l_{co}^{\eta})^{\frac{\sigma_A - 1}{\sigma_A}}$ and $L_n = (\phi_{ny} l_{ny}^{\eta} + \phi_{no} l_{no}^{\eta})^{\frac{\sigma_A}{\sigma_A - 1}}$ are the aggregate supply of college-educated and non-college workers across the two cohorts; ϕ_{je} denotes productivity of cohort j of educational level e; A_e denotes productivity of educational level e. Elasticity of substitution between educational level is denoted by σ_E and elasticity of substitution between cohort is denoted by σ_A.

Assuming that workers are paid at their marginal productivity, the college wage premium can be written as:[13]

$$\log \frac{w_{cj}}{w_{nj}} = \log \frac{A_c}{A_n} - \left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right). \log \frac{L_c}{L_n} - \frac{1}{\sigma_A} \log \frac{l_{cy}}{l_{ny}} \tag{2}$$

where the first three terms capture the general equilibrium effects on both young and s due to skill-biased technological change and a shift in the aggregate skill distribution, while the last term captures the additional effect on the younger cohort due to cohort-specific skill distribution shift.

In the short run, firms cannot adjust their production technology, so the relative productivity term, $\log \frac{A_c}{A_n}$, which reflects the type of technology firms use, is not dependent on the share of college-educated workers in the labor market. Therefore, the increase in the relative supply would lower the college wage premium, which is captured by the last two terms in Equation 2.

In the long run equilibrium, $\log \frac{A_c}{A_n}$ is dependent on the relative supply of college-educated workers in the given labor market. Firms can adopt production technology that either augments college-educated or non-college workers. Technology that augments college-educated workers, also known as skill-biased technology, is one that increases the marginal product of college-educated workers more than that of non-college workers. Technology that augments non-college workers, on the other hand, raises the marginal product of non-college workers more.

To understand how the relative supply affects firm productivity, suppose that the relative productivity $\log \frac{A_c}{A_n}$ is driven by the type of technology θ^k that firms choose to adopt. Firms choose to adopt technology θ^k over $\theta^{k'}$ when $Y(L_c, L_n, \theta^k) - c(\theta^k) > Y(L_c, L_n, \theta^{k'}) - c(\theta^{k'})$ where Y is the production function and $c(\theta)$ is the cost of adopting a technology.

[13] This is the canonical model of skill differentials (Katz and Murphy, 1992; Card and Lemieux, 2001).

It can be shown that firms will switch to skill-biased technology as the quantity of college-educated workers exceeds a certain threshold (Acemoglu, 2007; Carneiro et al., 2023).[14] As a result, we can write the relative productivity as a function of the relative supply of college-educated workers $\log \frac{A_c}{A_n} = f(\log \frac{L_c}{L_n})$. As the expansion increases the relative supply $\log \frac{L_c}{L_n}$, firms are more likely to adopt skill-biased technologies, raising the relative productivity $\log \frac{A_c}{A_n}$ of college-educated workers.

However, it is important to note that the technological adoption is a long-term equilibrium outcome, since the share of college-educated workers has to reach a certain threshold before firms find it worthwhile to invest in a new technology. In other words, we would expect the labor market to be affected first as the supply of college-educated workers in the younger cohort increases due to the expansion. Specifically, it reduces the college wage premium via a supply effect. In the long run, as firms adopt new technology, the relative productivity would increase and offset the negative wage effect.

This is similar to the technology-skill mismatch framework (Okoye, 2016; Acemoglu and Zilibotti, 2001)

Given the implications of this model, we structure the rest of the analysis as follows. We first analyze the effects on the labor market at the province level. We allow the treatment effects to vary between the short run (4 to 8 years after the expansion) and the long run (9 to 12 years after the expansion). We specifically study how the share as well as the relative supply of college-educated workers evolve in the short and long run for both cohorts and, in turn, how college wage premium changes in response.

Using Equation 2, we can further recover the structural parameters in this model such as

[14] Suppose there are two technologies that are skill-biased and labor-biased, denoted as θ^c and θ^n, and all firms start with the labor-biased technology θ^n. As the economy has more college-educated workers, the marginal product of college-educated workers under the skill-biased technology becomes higher than that under the labor-biased technology, i.e., $\frac{\partial Y}{\partial L_c}\Big|_{\theta^c} > \frac{\partial Y}{\partial L_c}\Big|_{\theta^n}$. Firms switch to skill-biased technology when the number of college-educated workers C exceeds a certain threshold C^* where firms are indifferent between the two technologies.

the elasticity of substitution across educational level, σ_E as well as the elasticity of substitution across cohort, σ_A. Specifically, we can assume that the technological adoption process does not happen in the short run, which we confirm with firm-level data later on.[15] Thus, $\left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right)$ captures the effect on college wage premiums of *both* cohorts, while $-\frac{1}{\sigma_A}$ captures the additional effect on college wage premium of the younger cohort only.

Given these GE effects on the labor market, we then turn to examine the effects on firm-level productivity. Similarly, we allow the treatment effects to vary in the short and long run. While $\log \frac{A_c}{A_n}$ captures the total factor productivity, we also examine if the results are consistent with labor productivity, which is measured by value added per worker.

This model belongs to a class of endogenous technological adoption models such as Beaudry et al. (2010), Clemens et al. (2018), and Blundell et al. (2022). There are other frameworks to think about how firms adapt to a labor supply shock.[16] Early model of human capital externalities assumes that an increase in the stock of college-educated workers would make everyone more productive (Moretti, 2004a). Therefore, an increase of college-educated workers would make the labor force more productive, offsetting some of the negative effect on relative wage. The empirical predictions of this model and ours are essentially identical.

Second, Acemoglu (1998, 2002) develops a model of directed technical change, in which an increase of college-educated workers would create an incentive for R&D firms to create innovations that are biased towards college-educated workers, raising the overall productivity level in sectors that use college-educated workers more. Given the lack of a formal R&D market, the directed technical change model is not applicable in our context.

Third, a class of capital-skill complementarity models argues that firms adopt skill-biased capital to replace unskilled labors in response to an increase of the relative supply of skilled

[15]This assumption is also commonly made in this literature. For example, see Card and Lemieux (2001) and Carneiro et al. (2023).

[16]See a more thorough discussion of these related frameworks in Lewis (2013).

labors because of the complementarity. As a result, productivity can also rise in such models. This alternative model is entirely possible in our context, given that other studies have found that firms adopt skill-biased capital in response to an increase in supply of skilled labor (Che and Zhang, 2018; Khanna, 2023).

There is a longstanding debate about whether firms adjust along the capital or production technique dimension in response to a change in skilled or unskilled labor (see Lewis (2011, 2013)). Our firm-level data allows us to directly test whether the capital skill complementarity model or the endogenous technological change model is correct by examining the effects on capital per worker and capital intensity.

5.2 Empirical Strategy for Market-Level Analysis

We can apply the stacked regression approach to study the effects of the expansion on firm-level productivity as well as market-level supply and college wage premium. Let $Y_{p,t}$ denote firm-level outcome in province p in year t and $Post_t$ indicates whether t is after the treatment starts. Similarly, we construct a subdataset for each group of provinces with the same treatment year g with a clean control group. The group-specific TWFE model is:

$$Y_{p,t}^{G} = \beta^{G}.(T_{p}^{G} \times Post_{t}^{G}) + \theta_{p} + \kappa_{t} + \epsilon_{p,t}$$

where β^G captures the province-level treatment effect for group G, while θ_p and κ_t capture province and year fixed effects. The analogous stacked regression at the market level is:

$$Y_{p,t,s} = \beta.(T_{p,s} \times Post_{t,s}) + \theta_{p,s} + \kappa_{t,s} + \epsilon_{p,t,s} \tag{3}$$

where β is the weighted average of all β^G. The main difference between the cohort-based DiD model and this model is that the first one uses variation across birth cohorts while the second one uses variation across survey years.

The advantage of the stacked regression approach, relative to other DiD staggered timing estimators, is that we can naturally extend our setup to a triple-differences model. Given our theoretical predictions, we can estimate the elasticity of substitution between college-educated and non-college workers via the difference in treatment effects between the two cohorts. Therefore, we estimate the following triple differences model:

$$Y_{p,t,s,j} = \delta.(T_{p,s} \times Post_{t,s} \times Young_{j,s}) + \theta_{p,j,s} + \kappa_{t,j,s} + \gamma_{p,t,s} + \epsilon_{p,t,s} \tag{4}$$

where $\theta_{p,j,s}$, $\kappa_{t,j,s}$, and $\gamma_{p,t,s}$ denote province-by-cohort-by-subdataset, year-by-cohort-by-subdataset, and province-by-year-by-subdataset fixed effects.

5.3 Results for the Effects on the Labor Market and Firms

5.3.1 Market-Level Effects

In Table 5, we report the results for the effects of the expansion on several province-level outcomes by cohorts. First, we examine the effect on the share of college-educated workers at the province level in column (1). In the first four to eight years post-expansion, the share of college-educated workers among the younger cohort raised by 1.3 percentage points. After 9 years post-expansion, this share increased slightly by 1.6 percentage points. Both short-term and long-term results are statistically significant. In contrast, the saw effects close to zero (0.4 percentage points) and statistically insignificant. Interestingly, even after 8 years, the effect remains very small. Figure 6(a) suggest that these effects are not driven by pre-trends.

The expansion also has a positive effect on employment rate among the college-educated workers, although the effect is larger and statistically significant for the younger cohort. Interestingly, the long-run employment effect is twice as big as the short-run effect for college-educated workers. This is suggestive that firms adjust their skill-biased technology in the long run and, thus, hire more college-educated workers. The expansion also has positive employment effects on non-college workers in both cohorts. The younger cohort, again, appears to experience larger employment effects than the , although the short-run and long-run employment effects are relatively similar. It is, however, important to note that some of these effects might be driven by pre-treatment trends, as suggested in Figure 6(b) and (c).

The expansion have negative but insignificant wage effect among college-educated workers in both cohorts. The long-run wage effects are smaller than the short-run effects although they are mostly statistically insignificant. In contrast, the wage effects on non-college workers are both positive and statistically significant. Most surprisingly, the long-term effects are much larger and statistically significant for both cohorts. These results point towards a higher demand for non-college workers in both cohorts in the long run. The stark contrast between the college-educated and non-college workers can also be observed in Figure 6(d) and (e).

We now turn to the effects on the relative supply and college wage premium in the last two columns. As expected, the expansion raises the relative supply of college-educated workers among the younger cohort in both short and long run. In the short run, the relative supply increased by 13.3 percentage points among the younger cohort. As expected, college premium declined by 11.7 percentage points due to the supply effect. More surprisingly, college premium of the younger cohort decreased almost twice as much in the long run as in the short run; the long-run effect of the expansion is 20.6%, while the relative supply did not appear to change. For the , these patterns are similar. The expansion has very small effects on the relative supply of skilled labors in both short and long run. College premium among the also suffered negative effects in both short and long run; the magnitudes of these effects are slightly higher than those among the younger cohort. The event study in Figure 7 paints a similar conclusion. Pre-trends do not appear to drive the effects on relative supply and premium.

In summary, the short-run effects are as expected from the theoretical model. The expansion shifts the relative supply of college-educated workers in the younger cohort. As a result, it reduces the college wage premium for both cohorts, suggesting that the two cohorts are close substitutes in production. In the long run, the effects on college wage premium are more negative in the long run, driven by substantial increases in the monthly wage of non-college workers. These long-run results suggest that the expansion creates a larger demand for unskilled labor in the long run. We also find that the expansion has a positive employment effect on non-college workers in both short and long run.

The larger negative effect on college premium in the long run is unexpected from both the theoretical model and also findings from the previous literature (Goldin and Katz, 2010; Blundell et al., 2022; Carneiro et al., 2023). That is, once firms can adjust their production technology in the long run, we would expect the negative effect on college wage premium would revert back as more firms adopt skill-biased technology and, thus, raise demand for college-educated workers. We come back to this after examining whether firms adopt better technology in the next part.

As discussed above, we can use these results to back out the relevant elasticities of substitution in Equation 2. Specifically, we first assume that technological adoption does not take place in the short run,[17] so the college wage premium effect on the in the short run, scaled by the respective relative supply effect, captures the effect of the aggregate skill distribution on college wage premium, i.e., $\left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right)$. The additional effect on the younger cohort (still in the short run) in the triple-differences model, scaled by the respective relative supply effect, captures the additional GE effect on the younger cohort (given that the two cohorts are imperfect substitutes), which is $\frac{1}{\sigma_A}$.

[17] This standard assumption is also made in other studies in this literature, such as Card and Lemieux (2001), Blundell et al. (2022), and Carneiro et al. (2023).

Given that the triple-differences estimate for college premium after 4 to 8 years is statistically indistinguishable from zero, we can conclude that young and old workers are perfect substitutes in production. Assuming that $\frac{1}{\sigma_A}$ is zero, the elasticity of substitution between college and non-college workers is 2.12. Our result lies well within the range of estimates for this parameter from the previous literature (Acemoglu and Autor, 2011).

5.3.2 Firm-Level Effects

We now turn to the results for the effects of the expansion on firm-level outcomes, which we report in Table 6. Same as before, we report separately the short-run and long-run effects. Our main outcome is total factor productivity (TFP), which we discuss the details of constructing in A. We also report the results for labor productivity, measured as log value added per worker, capital per capita, measured as log $\frac{assets}{labor}$, and capital intensity, measured as log $\frac{assets}{revenue}$.

Since different sectors can respond differently to the labor supply shock, we split our sample into service firms and manufacturing firms. We drop agricultural firms as well as firms in the energy, water, and waste, who are mostly controlled by the government. These firms tend to have different level of access to capital (Baccini et al., 2019) and thus, may respond differently than the rest of the sample.

First, we find that in the short run, service firms experience almost no effect on TFP (8 percentage points and not statistically significant). In the long run, the expansion increases their log TFP by 31.2%. This is consistent with our model that firms can only adjust their production technology in the long run in response to a change in the labor supply. This result is robust when we apply PSM to select the appropriate control group or when we allow for differential linear trends. We also find similar results when using labor productivity as an alternative outcome. In contrast, we find that the expansion has positive effects in both short and long run on manufacturing firms' TFP; the short-run effect is 12.4% while the long-run effect is 31.7%. These results are also robust with different specifications and different measures of productivity.

These results are consistent with our model of endogenous technological adoption. The increase in the number of college-educated workers raises the marginal product of these workers more than non-college workers under the skill-biased technology than labor-biased technology. Thus, when the number of skilled labor reaches a certain threshold, firms would find it worthwhile to adopt the skill-biased technology.

As discussed above, an alternative explanation for the rise of productivity is that firms adopt more capital as the number of college-educated workers goes up due to the complementarity nature of capital and skill; this increase in capital adoption would also lead to higher productivity (Lewis, 2013; Khanna, 2023). We can test this hypothesis by examining the effects on capital per capita and capital intensity. Interestingly, we find that not only capital did not rise, it appears to even decrease in the long run. These results suggest that firms were switching to a different model of production in response to the expansion. In other words, the capital-skill complementarity framework does not appear to apply in our context.

Two results stand out from our earlier discussion. We find that the expansion has a positive long-run effect on total factor productivity and a negative long-run effect on college wage premium. Generally, we would expect that an increase in total factor productivity would mitigate the negative effect on college wage premium in the long run if the technological change is biased towards skilled labor. That is, if technological change increases the marginal productivity of college-educated workers (and thus, their wage), then in the long run, skill-biased technological change would mitigate the supply effect in Equation 2.

There are two possible explanations for this puzzling result. First, it is possible that we are only observing the medium-term effects in this study, so firms may not have time to adjust wage as quickly as they adjust to their new technology. Second, it is possible that the expansion has positive spillover effects on demand for non-college workers. As Mazzolari and Ragusa (2013) and Liu and Yang (2021) note, high-skill labors may have demand for low-skill consumption such as domestic helpers or food services. The expansion might have raised the demand for such services, thus increasing non-college wage even further.

6 Partial equilibrium returns to higher education

The previous sections suggest that returns to higher education is subject to several general equilibrium effects. As discussed in Section **??**, the cohort-based DiD estimates cannot be used to recover the returns to higher education because they would be confounded by these GE effects. Yet Equation 5 also provides a way to recover the partial equilibrium returns to higher education.

Revisiting the equation, we note that we can, and have, identify the GE effects on wage of college-educated and non-college workers across both cohorts. It is also straightforward to obtain the shares of college-educated and non-college workers in both cohorts. Therefore, we can recover $\beta_{partial}$, which is the returns to college completion for an individual living in a control province.

$$W_{DiD} = \Delta l_{cy}.\beta_{partial} + (l_{cy,T=1}.\Delta w_{cy} + l_{ny,T=1}.\Delta w_{ny}) + (l_{co,T=1}.\Delta w_{co} + l_{no,T=1}.\Delta w_{no}) \quad (5)$$

where $l_{e,a}$ denotes the share of those with educational level $e = \{c, n\}$ (college and non-college) among cohort $a = \{y, o\}$ (young and old); $\Delta w_{e,a}$ are the GE effects on wage by educational level and age cohort; $\Delta l_{cy} = l_{cy,T=1} - l_{cy,T=0}$ is the difference in the share of college-educated workers in the younger cohort between the treatment and control provinces. The partial equilibrium returns is $\beta_{partial}$, which captures college completion returns for an individual in the control province.

Given that the GE effects vary in the short run and long run, we also estimate the partial equilibrium returns for both periods and report them in Table 7. The results suggest that higher education has a relatively high partial equilibrium rates of returns, between 235% and almost 400%, after accounting for the GE effects. The return using the long-run estimates is almost twice as smaller than the return using the short-run estimates; in other words, the return to higher education diminishes over time. It is also important to note that these estimates reflect the returns for those who "switched" to complete college because of the expansion. In other words, they may not be representative for other populations.

7 Conclusion

In the past three decades, Vietnam has experienced rapid economic growth due to both within-sectors productivity growth and structural transformation (McCaig and Pavcnik, 2013; McMillan et al., 2017; Liu et al., 2020). As noted by McMillan et al. (2017), structural transformation may lead to episodic growth but it requires more fundamental changes such as human capital investment and institutional changes to sustain economic growth. Decree 121/2007 is an effort to push for such fundamental changes, as it allowed Vietnam to establish over 100 new universities in a short period.

At the individual level, the expansion increases the probability of completing college by over 57%. It also raises their monthly wage by over 8%. Being exposed to the expansion also allows individual to work in skill-intensive sectors, better-paid jobs, and, thus, have higher wage. The expansion also affects firms through changes in local labor markets. As it raises the relative supply of college-educated workers, it increases the marginal product of college-educated workers in more productive technologies, thus inducing firms to adopt such technologies and raising firm productivity. Surprisingly, the expansion also raises non-college wage substantially in the long run, thus lowering the college wage premium even further.

These findings allow us to draw a number of policy implications for expanding access to higher education. First, expanding access to higher education does not appear to hurt college-educated workers, despite certain concerns about the quality of higher education and the lack of demand for college-educated workers in developing countries. These concerns were also apparent in Vietnam (The World Bank, 2020). Our findings indicate that exposure to the expansion improves individuals' job market outcomes in terms of employment and wage. At the labor market level, it does not hurt college-educated workers' wage. The returns to higher education is, in fact, very high even after accounting for the general equilibrium effects.

Second, higher education plays a significant role in fostering productivity growth in developing countries such as Vietnam. Our results indicate that expanding access to higher education contributes to economic growth via two channels. The expansion induces workers to move out of less productive sectors and into more productive sectors, speeding up the structural transformation process. More importantly, the expansion also forces firms to adopt better technology, raising their productivity. This finding addresses the concern that less developed countries may lack a formal market for innovation (Acemoglu, 1998, 2002) or an incentive to focus on innovative activities (Acemoglu et al., 2006; Vandenbussche et al., 2006; Aghion et al., 2009) and, thus,an expansion of high-skilled labors may not lead to more innovation activities.

Our study also indicates that the expansion benefits female workers more than male workers both in terms of college completion and wage. These results imply that the expansion also lowers the wage gap across gender, which is an important goal of economic development itself.

Our study provides an important case study on the effects of expanding access to higher education in developing countries; however, generalizing its results to other developing countries would require careful consideration because Vietnam may differ from other developing countries in many dimensions. Despite being a low-middle income country, Vietnam has been well-recognized for its success in basic education in terms of enrollment and international assessment scores (Dang and Glewwe, 2018; Dang et al., 2021). Such a factor may contribute to the effectiveness of the expansion. Since we focus on new universities in provinces that never had a university before, we mostly capture the effects of public universities.[18] Therefore, our results cannot be generalized to the effects of private universities.

Furthermore, it is important to note that Vietnam has enjoyed rapid growth in both GDP per capita as well as economic complexity during the expansion and, thus, it has also enjoyed a relatively high demand for high-skilled labor (Patrinos et al., 2018). This factor likely plays a crucial role in explaining the positive effects of the expansion on labor market outcomes. Thus, we can also expect that our results can be generalized to countries with a similar growth experience such as Bangladesh, Cambodia, Ethiopia, and Philippines (see Figure A8).

[18]Private universities tend to be located in large cities where there have already been existing universities before the reform.

References

Acemoglu, D. (1998). Why do new technologies complement skills? directed technical change and wage inequality. The Review of Economic Studies 113(4), 1055–1089.

Acemoglu, D. (2002). Directed technical change. The Review of Economic Studies 69(4), 781–809.

Acemoglu, D. (2007). Equilibrium bias of technology. Econometrica 75(5), 1371–1409.

Acemoglu, D., P. Aghion, and F. Zilibotti (2006). Distance to frontier, selection, and economic growth. Journal of the European Economic association 4(1), 37–74.

Acemoglu, D. and D. Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. In Handbook of Labor Economics, Volume 4, pp. 1043–1171. Elsevier.

Acemoglu, D. and F. Zilibotti (2001). Productivity differences. The Quarterly Journal of Economics 116(2), 563–606.

Ackerberg, D. A., K. Caves, and G. Frazer (2015). Identification properties of recent production function estimators. Econometrica 83(6), 2411–2451.

Aghion, P., L. Boustan, C. Hoxby, and J. Vandenbussche (2009). The causal impact of education on economic growth: evidence from US. Brookings papers on economic activity 1(1), 1–73.

Akresh, R., D. Halim, and M. Kleemans (2018). Long-term and intergenerational effects of education: Evidence from school construction in Indonesia. Technical report, National Bureau of Economic Research.

Andrews, M. (2020). How do institutions of higher education affect local invention? evidence from the establishment of US colleges. Evidence from the Establishment of US Colleges (March 28, 2020).

Arteaga, C. (2018). The effect of human capital on earnings: Evidence from a reform at colombia's top university. Journal of Public Economics 157, 212–225.

Athey, S. and G. W. Imbens (2006). Identification and inference in nonlinear difference-in-differences models. Econometrica 74(2), 431–497.

Baccini, L., G. Impullitti, and E. J. Malesky (2019). Globalization and state capitalism: Assessing vietnam's accession to the wto. Journal of International Economics 119, 75–92.

Barrera-Osorio, F. and H. Bayona-Rodríguez (2019). Signaling or better human capital: Evidence from colombia. Economics of Education Review 70, 20–34.

Beaudry, P., M. Doms, and E. Lewis (2010). Should the personal computer be considered a technological revolution? Evidence from US metropolitan areas. Journal of Political Economy 118(5), 988–1036.

Bilinski, A. and L. A. Hatfield (2018). Nothing to see here? non-inferiority approaches to parallel trends and other model assumptions. arXiv preprint arXiv:1805.03273.

Blundell, R., D. A. Green, and W. Jin (2022). The uk as a technological follower: Higher education expansion and the college wage premium. The Review of Economic Studies 89(1), 142–180.

Borusyak, K., X. Jaravel, and J. Spiess (2021). Revisiting event study designs: Robust and efficient estimation. arXiv preprint arXiv:2108.12419.

Buera, F. J. and J. P. Kaboski (2012). The rise of the service economy. American Economic Review 102(6), 2540–2569.

Buera, F. J., J. P. Kaboski, R. Rogerson, and J. I. Vizcaino (2022). Skill-biased structural change. The Review of Economic Studies 89(2), 592–625.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. Journal of Econometrics 225(2), 200–230.

Card, D. and T. Lemieux (2001). Can falling supply explain the rising return to college for younger men? a cohort-based analysis. The Review of Economic Studies 116(2), 705–746.

Carneiro, P., K. Liu, and K. G. Salvanes (2023). The supply of skill and endogenous technical change: evidence from a college expansion reform. Journal of the European Economic Association 21(1), 48–92.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The effect of minimum wages on low-wage jobs. The Quarterly Journal of Economics 134(3), 1405–1454.

Che, Y. and L. Zhang (2018). Human capital, technology adoption and firm performance: Impacts of China's higher education expansion in the late 1990s. The Economic Journal 128(614), 2282–2320.

Chen, X., G. Pei, Z. Song, and F. Zilibotti (2023). Tertiarization like china. Annual Review of Economics 15.

Clemens, M. A., E. G. Lewis, and H. M. Postel (2018). Immigration restrictions as active labor market policy: Evidence from the Mexican Bracero exclusion. American Economic Review 108(6), 1468–87.

Coxhead, I. and R. Shrestha (2017). Globalization and school–work choices in an emerging economy: Vietnam. Asian Economic Papers 16(2), 28–45.

Dang, H.-A., P. Glewwe, K. Vu, and J. Lee (2021). What explains vietnam's exceptional performance in education relative to other countries? analysis of the 2012 and 2015 pisa data.

Dang, H.-A. H. and P. W. Glewwe (2018). Well begun, but aiming higher: A review of Vietnam's education trends in the past 20 years and emerging challenges. The Journal of Development Studies 54(7), 1171–1195.

Daw, J. R. and L. A. Hatfield (2018). Matching and regression to the mean in difference-in-differences analysis. Health Services Research 53(6), 4138–4156.

De Chaisemartin, C. and X. d'Haultfoeuille (2020). Two-way fixed effects estimators with heterogeneous treatment effects. American Economic Review 110(9), 2964–96.

Doan, T., Q. Le, and T. Q. Tran (2018). Lost in transition? Declining returns to education in Vietnam. The European Journal of Development Research 30(2), 195–216.

Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. American Economic Review 91(4), 795–813.

Elsayed, A. and A. Shirshikova (2023). The women-empowering effect of higher education. Journal of Development Economics, 103101.

Fan, T., M. Peters, and F. Zilibotti (2023). Growing like india—the unequal effects of service-led growth. Econometrica 91(4), 1457–1494.

Feng, S. and X. Xia (2022). Heterogeneous firm responses to increases in high-skilled workers: Evidence from China's college enrollment expansion. China Economic Review 73, 101791.

Goldin, C. and L. F. Katz (2010). The race between education and technology. Harvard University Press.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of Econometrics 225(2), 254–277.

Hausman, N. (2020). University innovation and local economic growth. The Review of Economics and Statistics, 1–46.

Hausmann, R., C. A. Hidalgo, S. Bustos, M. Coscia, and A. Simoes (2014). The Atlas of economic complexity: Mapping paths to prosperity. Mit Press.

Hendricks, L. and T. Schoellman (2023). Skilled labor productivity and cross-country income differences. American Economic Journal: Macroeconomics 15(1), 240–268.

Huang, B., M. Tani, Y. Wei, and Y. Zhu (2022). Returns to education in China: Evidence from the great higher education expansion. China Economic Review, 101804.

Imbens, G. W. and J. M. Wooldridge (2009). Recent developments in the econometrics of program evaluation. Journal of Economic Literature 47(1), 5–86.

Katz, L. F. and K. M. Murphy (1992). Changes in relative wages, 1963–1987: supply and demand factors. The Review of Economic Studies 107(1), 35–78.

Khanna, G. (2023). Large-scale education reform in general equilibrium: Regression discontinuity evidence from India. Journal of Political Economy 131(2), 000–000.

Levinsohn, J. and A. Petrin (2003). Estimating production functions using inputs to control for unobservables. The Review of Economic Studies 70(2), 317–341.

Lewis, E. (2011). Immigration, skill mix, and capital skill complementarity. The Review of Economic Studies 126(2), 1029–1069.

Lewis, E. (2013). Immigration and production technology. Annual Review of Economics 5(1), 165–191.

Li, H., P. Loyalka, S. Rozelle, and B. Wu (2017). Human capital and China's future growth. Journal of Economic Perspectives 31(1), 25–48.

Li, H., Y. Ma, L. Meng, X. Qiao, and X. Shi (2017). Skill complementarities and returns to higher education: Evidence from college enrollment expansion in China. China Economic Review 46, 10–26.

Li, S., J. Whalley, and C. Xing (2014). China's higher education expansion and unemployment of college graduates. China Economic Review 30, 567–582.

Liu, S. (2015). Spillovers from universities: Evidence from the land-grant program. Journal of Urban Economics 87, 25–41.

Liu, S. and X. Yang (2021). Human capital externalities or consumption spillovers? the effect of high-skill human capital across low-skill labor markets. Regional Science and Urban Economics 87, 103620.

Liu, Y., C. B. Barrett, T. Pham, and W. Violette (2020). The intertemporal evolution of agriculture and labor over a rapid structural transformation: Lessons from vietnam. Food Policy 94, 101913.

Martellini, P., T. Schoellman, and J. Sockin (2022). The global distribution of college graduate quality. Federal Reserve Bank of Minneapolis Working Paper 791.

Mazzolari, F. and G. Ragusa (2013). Spillovers from high-skill consumption to low-skill labor markets. Review of Economics and Statistics 95(1), 74–86.

McCaig, B. and N. Pavcnik (2013). Moving out of agriculture: structural change in vietnam. Technical report, National Bureau of Economic Research.

McGuinness, S., E. Kelly, T. T. P. Pham, T. T. T. Ha, and A. Whelan (2021). Returns to education in Vietnam: A changing landscape. World Development 138, 105205.

McMillan, M., D. Rodrik, and C. Sepulveda (2017). Structural change, fundamentals and growth: A framework and case studies. Technical report, National Bureau of Economic Research.

Melly, B. and G. Santangelo (2015). The changes-in-changes model with covariates. Universität Bern, Bern.

Moretti, E. (2004a). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. Journal of Econometrics 121(1-2), 175–212.

Moretti, E. (2004b). Workers' education, spillovers, and productivity: evidence from plant-level production functions. American Economic Review 94(3), 656–690.

Newman, C., J. Rand, T. Talbot, and F. Tarp (2015). Technology transfers, foreign investment and productivity spillovers. European Economic Review 76, 168–187.

Okoye, D. (2016). Appropriate technology and income differences. International Economic Review 57(3), 955–996.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment. Econometrica 64(6), 1263–1297.

Parajuli, D., D. K. Vo, J. Salmi, and N. T. A. Tran (2020). Improving the performance of higher education in vietnam: Strategic priorities and policy options.

Patrinos, H. A., P. V. Thang, and N. D. Thanh (2018). The economic case for education in Vietnam. World Bank Policy Research Working Paper (8679).

Phan, D. and I. Coxhead (2013). Long-run costs of piecemeal reform: wage inequality and returns to education in Vietnam. Journal of Comparative Economics 41(4), 1106–1122.

Porzio, T., F. Rossi, and G. Santangelo (2022). The human side of structural transformation. American Economic Review 112(8), 2774–2814.

Rodrik, D. (2016). Premature deindustrialization. Journal of economic growth 21, 1–33.

Russell, L. C., L. Yu, and M. J. Andrews (2022). Higher education and local educational attainment: Evidence from the establishment of US colleges. The Review of Economics and Statistics, 1–32.

Sánchez, A. and A. Singh (2018). Accessing higher education in developing countries: Panel data analysis from India, Peru, and Vietnam. World Development 109, 261–278.

Schoellman, T. (2012). Education quality and development accounting. The Review of Economic Studies 79(1), 388–417.

The World Bank (2020). Improving the performance of higher education in vietnam: Strategic priorities and policy options.

The World Bank (2022). Gross enrolment ratio for tertiary education, both sexes. data retrieved from World Development Indicators.

Valero, A. (2021). Education and economic growth.

Vandenbussche, J., P. Aghion, and C. Meghir (2006). Growth, distance to frontier and composition of human capital. Journal of Economic Growth 11(2), 97–127.

Vu, L. H. and A. T. Nguyen (2018). Analysis of access and equity in the vietnamese higher education system. VNU Journal of Science: Policy and Management Studies 34(4), 65–80.

Xing, C., P. Yang, and Z. Li (2018). The medium-run effect of China's higher education expansion on the unemployment of college graduates. China Economic Review 51, 181–193.

TABLES

Table 1: Balance table of treatment status at the province level

	Control (N = 14)	Treatment (N = 21)		Already treated (N = 28)	
	Mean	Mean	Difference	Mean	Difference
% college graduates	0.033	0.042	0.009	0.052	0.019*
			(0.011)		(0.011)
% high school graduates	0.243	0.256	0.013	0.302	0.060**
			(0.025)		(0.024)
% college enrolment	0.190	0.295	0.105***	0.301	0.111***
			(0.033)		(0.031)
% self employed	0.889	0.858	-0.031	0.821	-0.068**
			(0.028)		(0.026)
% employed	0.111	0.142	0.031	0.179	0.068**
			(0.028)		(0.026)
% agricultural worker	0.748	0.603	-0.146***	0.498	-0.250***
			(0.052)		(0.049)
% manufacturing worker	0.076	0.159	0.083***	0.212	0.136***
			(0.026)		(0.025)
% service worker	0.146	0.211	0.066**	0.264	0.118***
			(0.032)		(0.030)
Log income per capita	8.555	8.755	0.200**	8.881	0.326***
			(0.094)		(0.089)
Urban	0.154	0.213	0.059	0.283	0.128**
			(0.054)		(0.051)
% in poverty	0.182	0.134	-0.047**	0.108	-0.074***
			(0.021)		(0.020)
% age 0-5	0.238	0.214	-0.024*	0.201	-0.037***
			(0.014)		(0.013)
% age 6-18	0.176	0.180	0.005	0.170	-0.005
			(0.006)		(0.006)
TFP	11.628	11.751	0.124	11.702	0.075
			(0.225)		(0.214)
Labor productivity	16.982	16.990	0.008	16.997	0.015
			(0.110)		(0.104)
Capital-labor ratio	19.255	19.230	-0.025	19.383	0.128
			(0.104)		(0.099)
Total N	64				

This table shows the means of pre-treatment, province-level characteristics by treatment status and the results from their balance tests. The *Mean* columns display the mean values of these covariates for each treatment group. The *Difference* columns show the results from estimating a regression with the characteristics as the dependent variable and the treatment status dummy variables as the independent variables. Data on the pre-treatment characteristics are aggregated from the 2004-2006 VHLSS data.

Table 2: Summary statistics by cohort and province

	Treatment			Control		
Age exposed	26-36	19-25	12-18	26-36	19-25	12-18
Age	38.602	29.756	24.555	38.017	29.236	24.404
	(3.916)	(3.052)	(2.064)	(3.875)	(2.999)	(2.050)
Female	0.509	0.503	0.488	0.504	0.499	0.483
	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)
Completed college or higher	0.102	0.123	0.108	0.120	0.115	0.077
	(0.302)	(0.329)	(0.311)	(0.325)	(0.319)	(0.266)
Employment	0.426	0.532	0.582	0.343	0.396	0.357
	(0.494)	(0.499)	(0.493)	(0.475)	(0.489)	(0.479)
Log monthly wage	3.315	3.273	3.205	3.160	3.062	2.944
	(0.706)	(0.630)	(0.592)	(0.870)	(0.827)	(0.842)
Ag. employment	0.375	0.343	0.330	0.517	0.505	0.569
	(0.484)	(0.475)	(0.470)	(0.500)	(0.500)	(0.495)
Manufact. employment	0.137	0.180	0.210	0.077	0.089	0.102
	(0.344)	(0.385)	(0.407)	(0.267)	(0.285)	(0.302)
Service employment	0.096	0.108	0.107	0.111	0.120	0.079
	(0.295)	(0.311)	(0.309)	(0.314)	(0.325)	(0.270)

Table 3: Difference-in-differences estimates for the effect of exposure to the higher education expansion on individual-level outcomes

	Complete college	Being employed	Log wage	Employment		
				Ag.	Manufact.	Service
	(1)	(2)	(3)	(4)	(5)	(6)
Specification 1: TWFE						
Exposed at 19 to 25 years old	0.022***	0.043***	0.035*	-0.014	0.029**	0.000
	(0.007)	(0.015)	(0.020)	(0.013)	(0.013)	(0.007)
Exposed at 14 to 18 years old	0.045***	0.131***	0.086**	-0.090***	0.046*	0.041***
	(0.011)	(0.033)	(0.037)	(0.028)	(0.024)	(0.009)
N	2909364	2731604	1821739	2731604	2731604	2731604
Specification 2: TWFE with PSM						
Exposed at 19 to 25 years old	0.026***	0.046**	0.040*	-0.020	0.022	0.006
	(0.009)	(0.020)	(0.022)	(0.015)	(0.017)	(0.006)
Exposed at 14 to 18 years old	0.044***	0.129***	0.082*	-0.095***	0.035	0.043***
	(0.014)	(0.044)	(0.045)	(0.034)	(0.032)	(0.009)
N	1912122	1784737	1244384	1784737	1784737	1784737
Specification 3: TWFE with differential linear trends						
Exposed at 19 to 25 years old	0.017***	0.009	0.012	-0.009	0.007*	-0.003
	(0.005)	(0.006)	(0.011)	(0.007)	(0.004)	(0.005)
Exposed at 14 to 18 years old	0.036***	0.074***	0.047	-0.082***	0.007	0.036***
	(0.012)	(0.017)	(0.028)	(0.016)	(0.008)	(0.011)
N	2909364	2731604	1821739	2731604	2731604	2731604
Specification 3: TWFE with non-migrant						
Exposed at 19 to 25 years old	0.022***	0.048***	0.041*	-0.021	0.033**	0.002
	(0.008)	(0.015)	(0.022)	(0.014)	(0.014)	(0.006)
Exposed at 14 to 18 years old	0.047***	0.137***	0.102**	-0.098***	0.046*	0.043***
	(0.013)	(0.036)	(0.043)	(0.033)	(0.025)	(0.010)
N	2082423	1965391	1526743	1965391	1965391	1965391
Specification 4: TWFE with province of origin						
Exposed at 19 to 25 years old	0.021***	0.043***	0.030	-0.014	0.030**	-0.002
	(0.007)	(0.016)	(0.022)	(0.015)	(0.013)	(0.007)
Exposed at 14 to 18 years old	0.047***	0.139***	0.089**	-0.097***	0.047**	0.040***
	(0.012)	(0.035)	(0.042)	(0.030)	(0.023)	(0.009)
N	1797157	1695068	1090311	1695068	1695068	1695068

Table reports DiD estimate for the effects on individual outcomes. For each group of provinces with the same treatment year in each survey year, we create a subdataset with those treatment provinces and provinces that were never treated as the control group. Subdatasets are stacked together and a TWFE model is estimated on this new dataset, controlling for province-by-subdataset and cohort-by-subdataset fixed effects. All models control for age, age squared, and gender. Specification 2 is a TWFE model with a matched control group based on PSM. Specification 3 is a TWFE model that allows treatment and control to have differential linear trends. Specification 4 is the same as the first but restricted to a non-migrant sample. Specification 5 is the same as the first but uses the province of origin to construct the treatment variable instead of province of current residence. All samples include individuals between age 22 and 55. All standard errors are clustered at the province level. Data is drawn from LFS 2010-2018.

Table 4: Summary statistics by year and district

	Control		Treatment	
	2011	2015-2019	2011	2015-2019
% of adults who complete college or higher				
	0.064	0.086	0.157	0.206
	(0.054)	(0.073)	(0.110)	(0.127)
% of adults who are employed				
	0.325	0.400	0.526	0.596
	(0.136)	(0.165)	(0.125)	(0.110)
% of non-college adults who are employed				
	0.285	0.353	0.454	0.513
	(0.129)	(0.164)	(0.106)	(0.104)
% of college-educated adults who are employed				
	0.924	0.891	0.903	0.893
	(0.095)	(0.130)	(0.076)	(0.058)
Log monthly wage				
	10.111	10.299	10.329	10.643
	(0.218)	(0.362)	(0.227)	(0.257)
Log monthly wage of non-college workers				
	10.034	10.245	10.199	10.568
	(0.223)	(0.386)	(0.200)	(0.238)
Log monthly wage of college-educated workers				
	10.399	10.664	10.577	10.848
	(0.226)	(0.254)	(0.235)	(0.259)
Skill premium				
	0.360	0.408	0.378	0.280
	(0.252)	(0.363)	(0.193)	(0.179)

Table 5: Difference-in-differences estimates for general equilibrium effects on labor market at province level

	College share	Employment: College	Employment: Non-college	Log wage: College	Log wage: Non-college	Relative supply	College premium
All cohorts							
4-8 years after	0.003	0.003	0.034***	-0.052***	0.078	0.054	-0.130*
	(0.004)	(0.006)	(0.011)	(0.018)	(0.066)	(0.050)	(0.067)
9+ years after	0.002	0.008	0.044**	-0.041**	0.237**	0.031	-0.278**
	(0.008)	(0.010)	(0.017)	(0.020)	(0.112)	(0.093)	(0.112)
Young cohort							
4-8 years after	0.013***	0.035***	0.052***	-0.040	0.077	0.133**	-0.117*
	(0.004)	(0.011)	(0.016)	(0.035)	(0.054)	(0.052)	(0.058)
9+ years after	0.016**	0.067***	0.064**	-0.003	0.204**	0.130	-0.206**
	(0.008)	(0.019)	(0.024)	(0.038)	(0.095)	(0.097)	(0.092)
Old cohort							
4-8 years after	0.004	0.008	0.039***	-0.044*	0.081	0.059	-0.125*
	(0.004)	(0.007)	(0.012)	(0.023)	(0.065)	(0.056)	(0.067)
9+ years after	0.004	0.016	0.049**	-0.026	0.239**	0.042	-0.266**
	(0.008)	(0.012)	(0.019)	(0.026)	(0.112)	(0.099)	(0.110)
Triple differences							
4-8 years after	0.012**	0.041***	0.020**	0.006	-0.006	0.111*	0.012
	(0.005)	(0.011)	(0.010)	(0.029)	(0.029)	(0.056)	(0.041)
9+ years after	0.018***	0.076***	0.021	0.035	-0.054	0.132	0.089*
	(0.006)	(0.017)	(0.013)	(0.036)	(0.047)	(0.078)	(0.050)
N	2716	2716	2716	2716	2716	2716	2716

This table reports DiD estimate for the effects of the expansion at the province level. For each treatment group of provinces with the same treatment year, we create a subdataset with those treatment provinces and provinces that were never treated as the control group. These subdatasets are stacked together and a stacked DiD model is estimated on this new dataset, controlling for province-by-subdataset and year-by-subdataset fixed effects. Only the DiD coefficient of the interaction term is reported. Young and Old cohorts are defined as those above or below 35 years old. All samples include individuals between age 22 and 55. Relative supply is defined as log of number of college-educated workers divided by number of non-college workers. Premium is defined as the log of wage ratio between the two workers. All standard errors are clustered at the province level.

Table 6: Difference-in-differences estimates for effects of the expansion on firm-level outcomes

	Service				Manufacturing			
	TFP	Labor productivity	Capital per capita	Capital intensity	TFP	Labor productivity	Capital per capita	Capital intensity
Specification 1: TWFE								
4-8 years after	0.088	0.018	0.039	0.017	0.124***	0.059*	0.017	-0.054
	(0.055)	(0.045)	(0.064)	(0.070)	(0.033)	(0.031)	(0.070)	(0.061)
9+ years after	0.312***	0.194***	-0.079	-0.195***	0.317***	0.125*	-0.195***	-0.186
	(0.103)	(0.059)	(0.103)	(0.072)	(0.057)	(0.066)	(0.072)	(0.095)
N	3093185	3101045	3565527	3161636	775841	777424	3161636	847216
Specification 2: TWFE with PSM								
4-8 years after	0.084	0.034	0.046	-0.023	0.113***	0.063*	-0.023	-0.003
	(0.057)	(0.060)	(0.075)	(0.090)	(0.035)	(0.035)	(0.090)	(0.058)
9+ years after	0.239**	0.164***	-0.059	-0.234***	0.305***	0.123*	-0.234***	-0.143
	(0.104)	(0.060)	(0.126)	(0.084)	(0.059)	(0.064)	(0.084)	(0.094)
N	2808882	2815861	3249452	2862601	717205	718601	2862601	784241
Specification 3: TWFE w/ linear differential trends								
4-8 years after	0.046	-0.030	0.059*	0.069	0.127**	0.146***	0.069	-0.034*
	(0.047)	(0.037)	(0.030)	(0.042)	(0.049)	(0.049)	(0.042)	(0.046)
9+ years after	0.245**	0.117	-0.047	-0.113	0.323***	0.269**	-0.113	-0.153
	(0.114)	(0.116)	(0.054)	(0.086)	(0.119)	(0.115)	(0.086)	(0.057)
N	3093185	3101045	3565527	3161636	775841	777424	3161636	847216

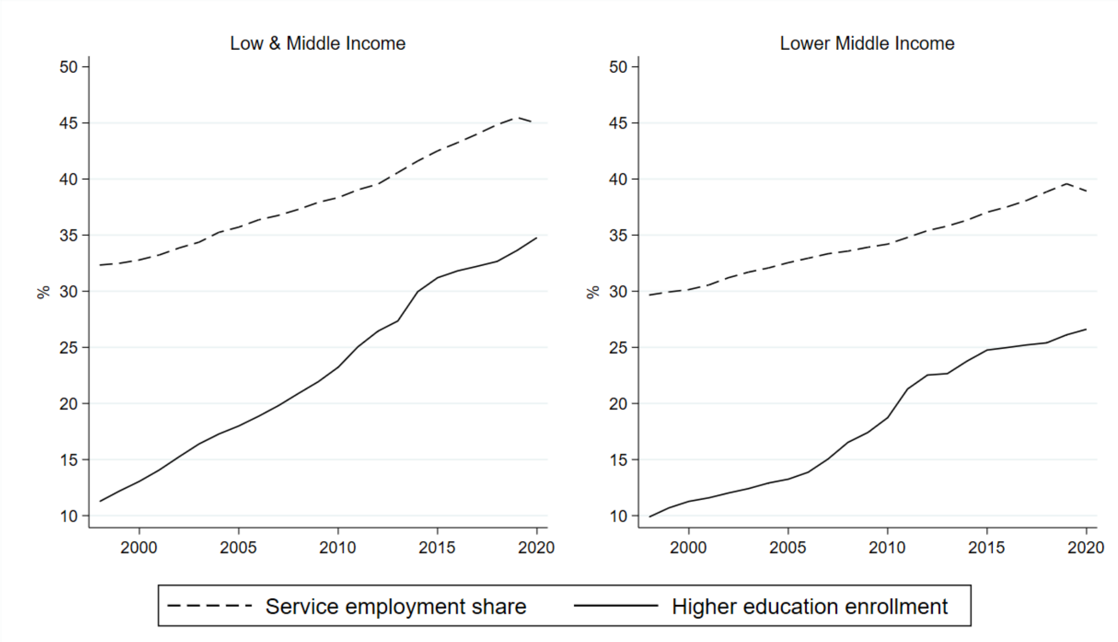
This table reports DiD estimate for the effects of the expansion at the individual level. For each treatment group of provinces with the same treatment year and for each survey year, we create a subdataset with those treatment provinces and provinces that were never treated as the control group. These subdatasets are stacked together and a stacked DiD model is estimated on this new dataset, controlling for province-by-subdataset and cohort-by-subdataset fixed effects. Only the DiD coefficient of the interaction term is reported. All models control for age, age squared, and gender. In specification 2, we alternatively use the TWFE model using a matched control group based on propensity score matching. In specification 3, we allow the treatment and control groups to have linear differential trends. All samples include individuals between age 22 and 55. All standard errors are clustered at the province level.

Table 7: Partial equilibrium returns to higher education in short run and long run

	Short run		Long run	
	College-educated	Non-college	College-educated	Non-college
Share of				
Old cohort:	0.086	0.914	0.092	0.908
	(0.028)	(0.028)	(0.023)	(0.023)
Young cohort:	0.095	0.905	0.097	0.903
	(0.029)	(0.029)	(0.019)	(0.019)
DiD Wage	0.087***		0.044*	
	(0.024)		(0.024)	
Partial equilibrium returns:	3.960		2.350	

FIGURES

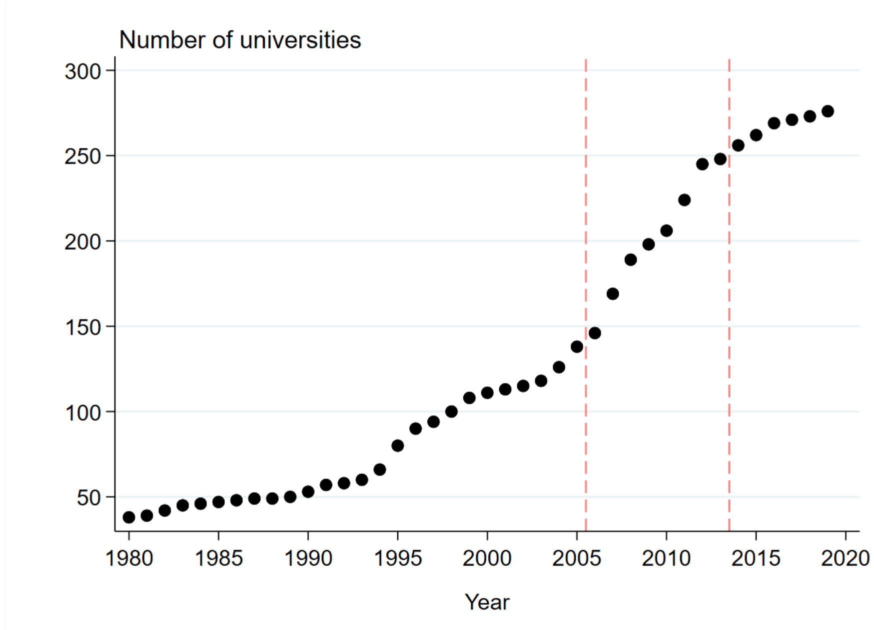
Figure 1: Service employment and higher education enrollment trends



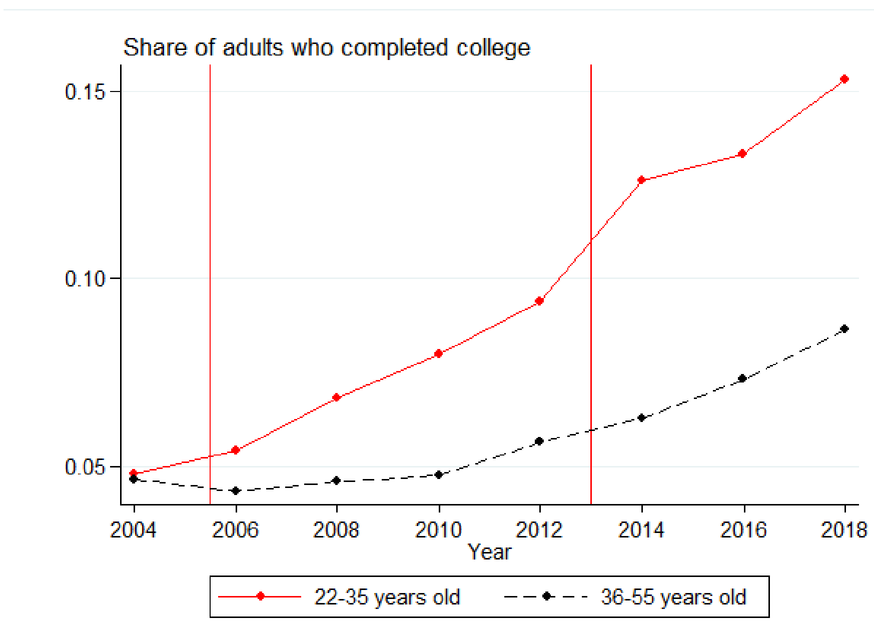
Source: World Development Indicators (The World Bank, 2022).

Figure 2: Number of universities and share of adults completing college education over time

(a) Number of universities by year

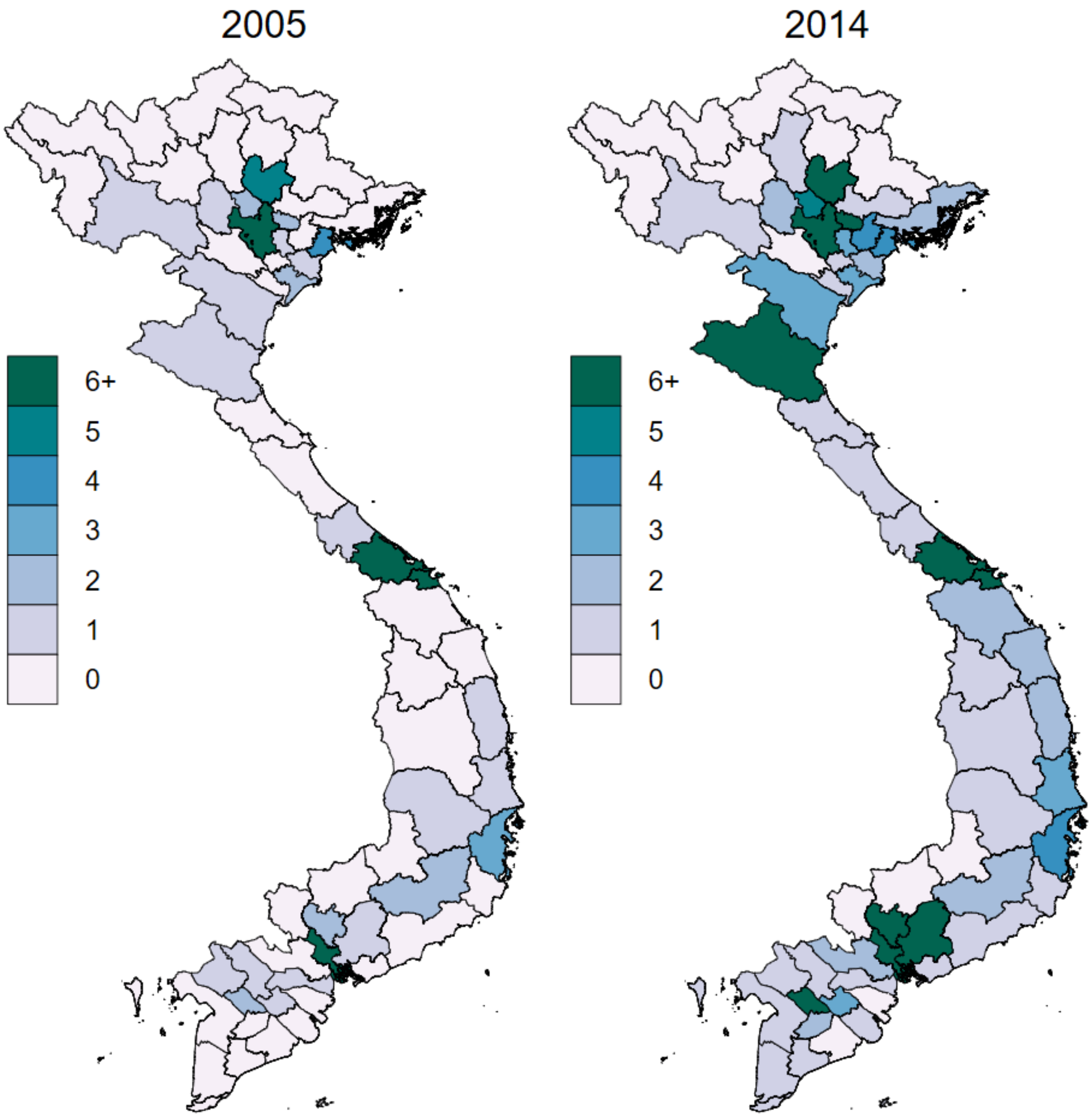


(b) Share of college-educated adults by age and year



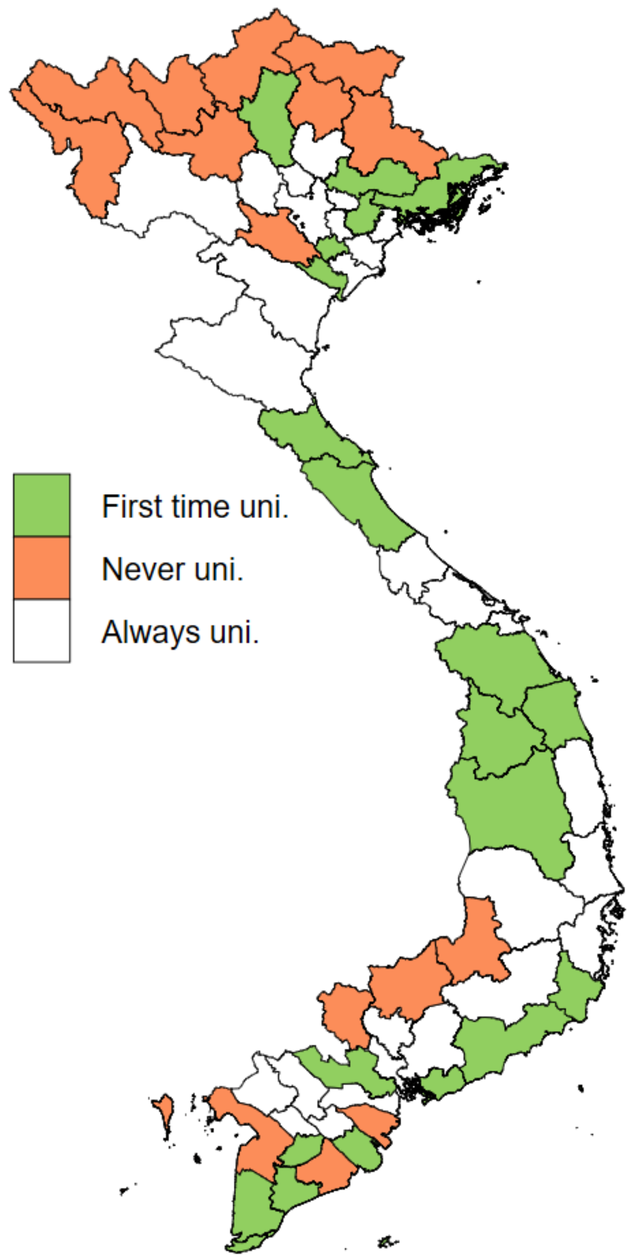
Note: Figure (a) shows the number of universities by year based on data collected from official documents. Figure (b) shows the share of adults between age 22-55 who completed college education based on data from the Vietnam Household Living Standard Survey for 2004-2018.

Figure 3: Number of universities by provinces in 2005 and 2014



Note: The graph shows the number of universities in each province in 2005 and 2015.

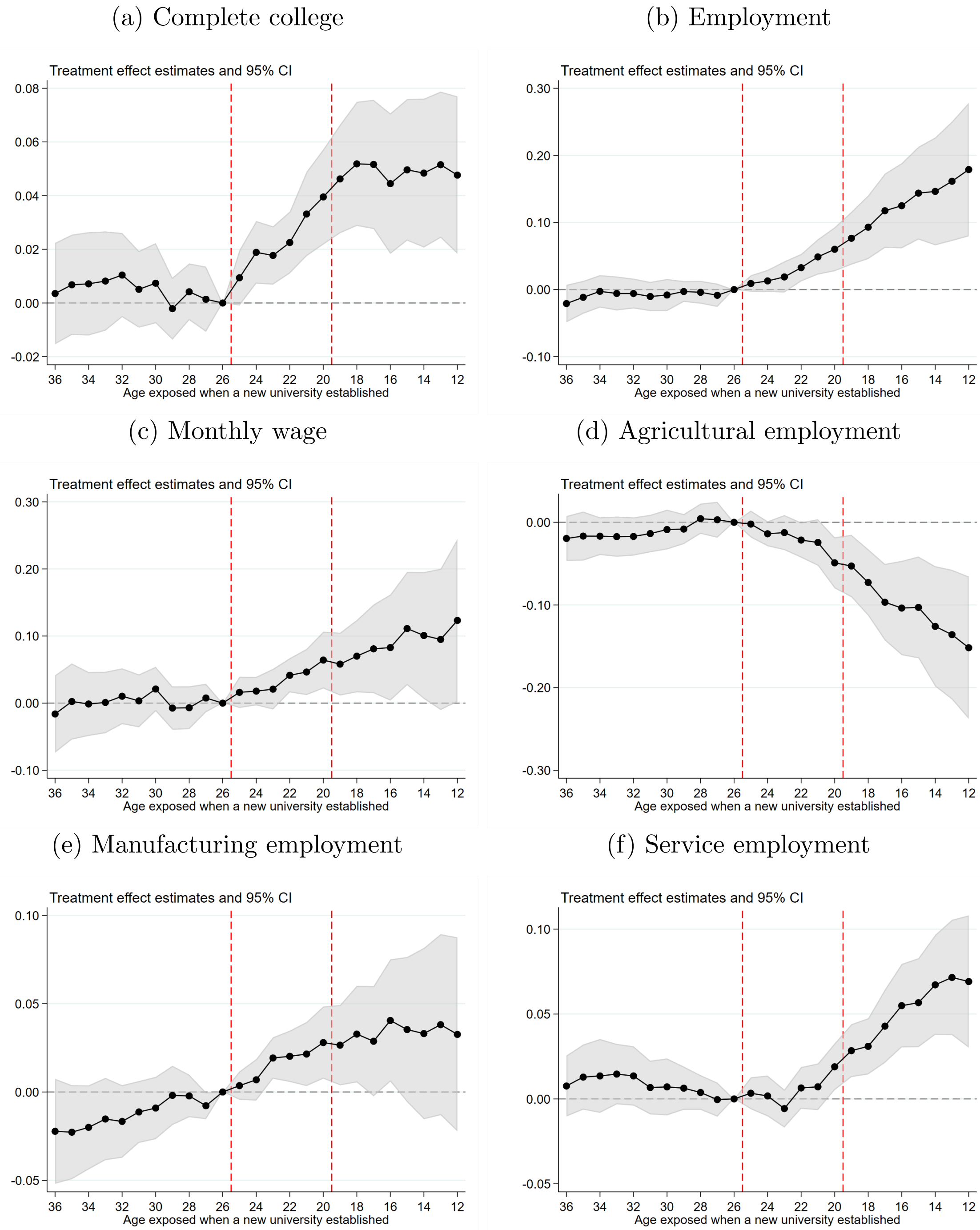
Figure 4: Treatment status of provinces



Source: Data on university location and opening date are collected by the authors from official documents.

Note: Treated provinces are those with a new university for the first time during the expansion. Never-treated provinces are those that never had a university before. Provinces that already had existing universities are not included in the cohort-based difference-in-differences analysis.

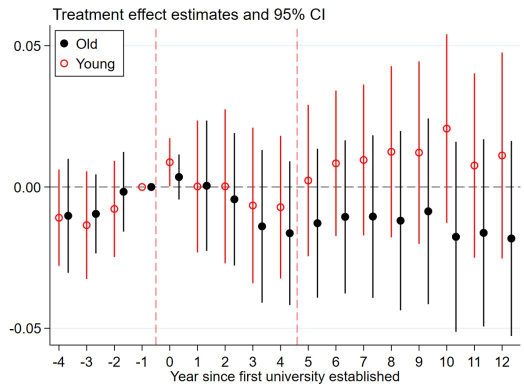
Figure 5: Event study estimation (cohort) for the effects of the higher education expansion



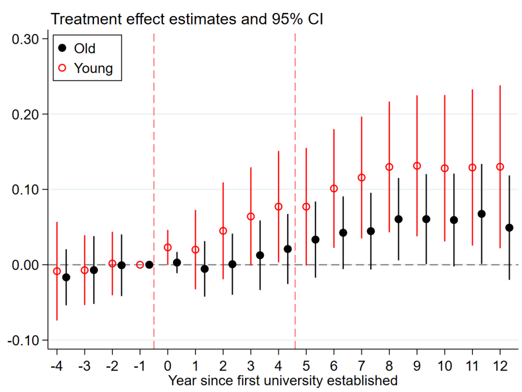
Note: The graphs display event study estimation for the effects of the higher education expansion on college completion and labor market outcomes. All models control for age, age squared, and gender. Standard errors are clustered at the province-level and 95% confidence intervals are displayed.

Figure 6: Event study results for the effects on province-level outcomes - by age cohort

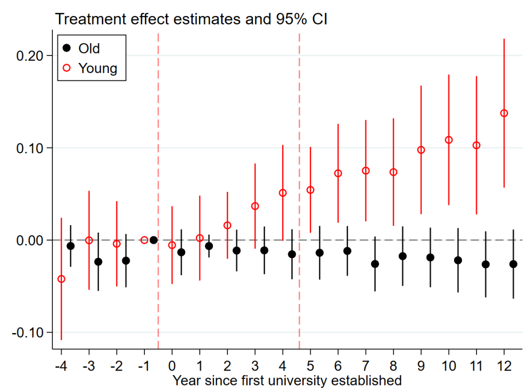
(a) Share of adults with college education



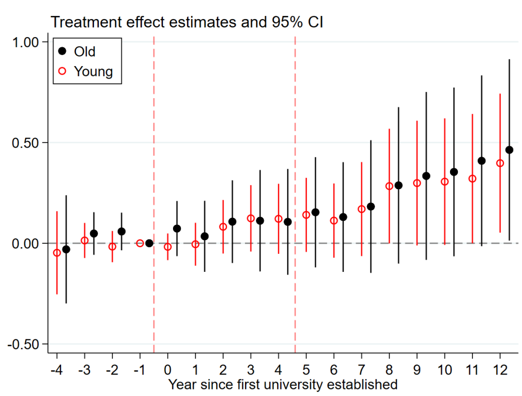
(b) Employment rate of non-college



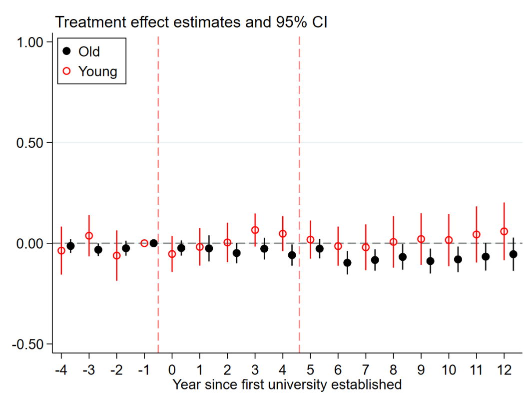
(c) Employment rate of college-educated



(d) Log wage of non-college



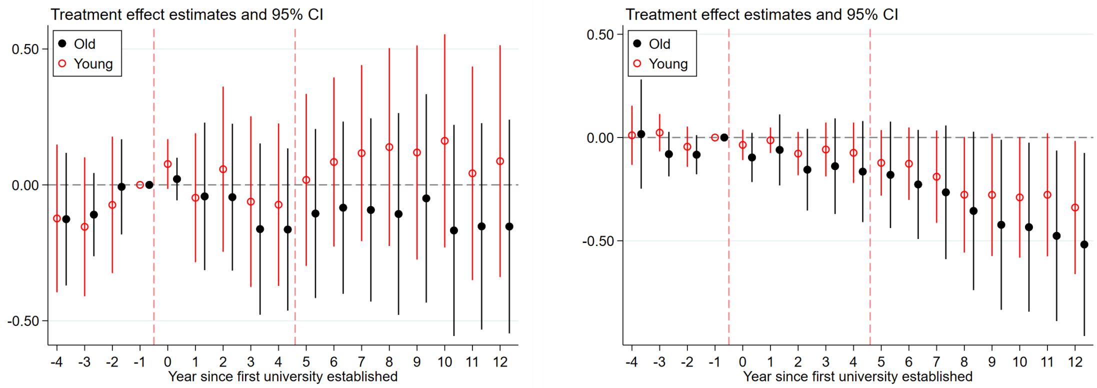
(e) Log wage of college-educated



Note: These graphs show the event-study estimation results for the effects province-level outcome. Each estimate shows the coefficient of the interaction term between year dummy and treatment status. All models control for province and year fixed effects. Standard errors are clustered at the province-level and 95% confidence intervals are displayed.

Figure 7: Event study results for the effects on labor market equilibrium

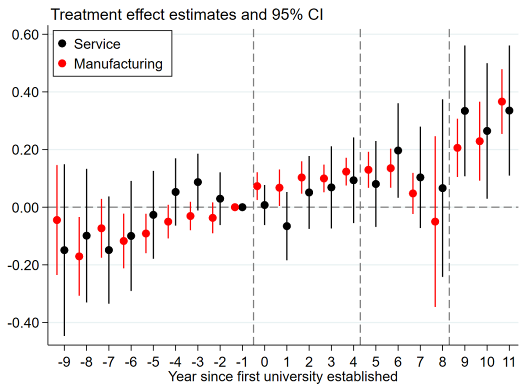
(a) Relative supply (b) College wage premium



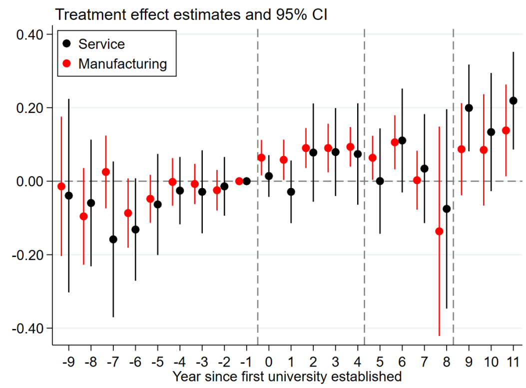
Note: These graphs show the event-study estimation results for the effects province-level outcome. Each estimate shows the coefficient of the interaction term between year dummy and treatment status. All models control for province and year fixed effects. Standard errors are clustered at the province-level and 95% confidence intervals are displayed.

Figure 8: Event study results for the effects on firm-level productivity

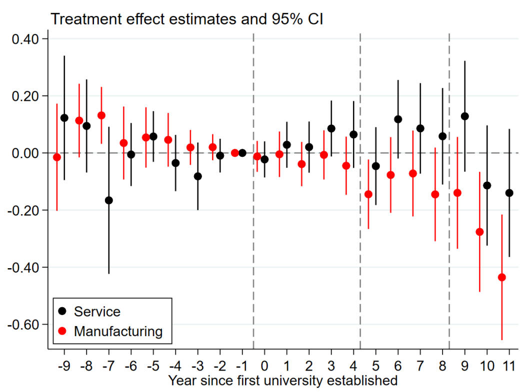
(a) Log TFP



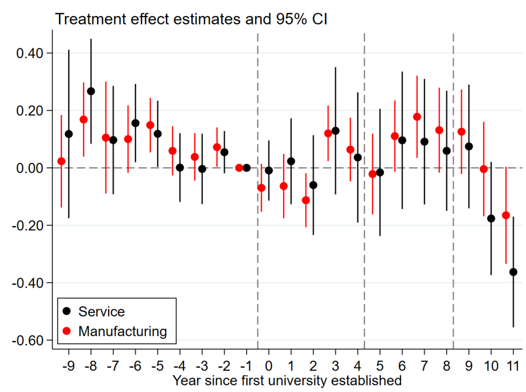
(b) Labor productivity



(c) Capital per capita



(d) Capital intensity

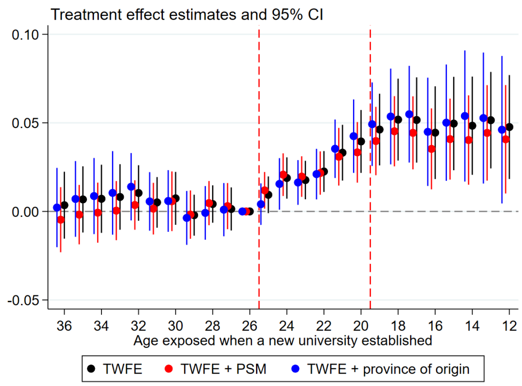


Note: These graphs show the event-study estimation results for the effects of the higher education expansion on firm-level outcome. Each estimate shows the coefficient of the interaction term between year dummy and treatment status. All models control for province and year fixed effects. Standard errors are clustered at the province-level and 95% confidence intervals are displayed. Capital intensity is measured by total value of capital divided by total revenue. Labor productivity is measured by value added per worker.

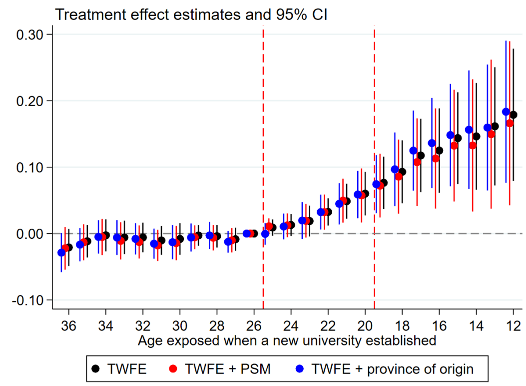
Appendix

Figure A1: Event study estimation (cohort) for the effects of the higher education expansion using different control groups

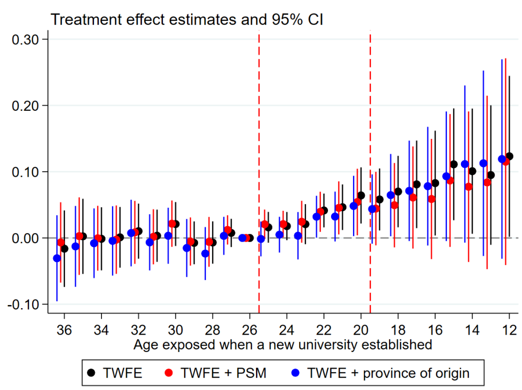
(a) Complete college



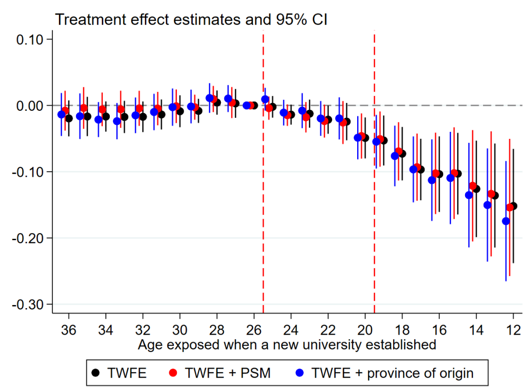
(b) Employment



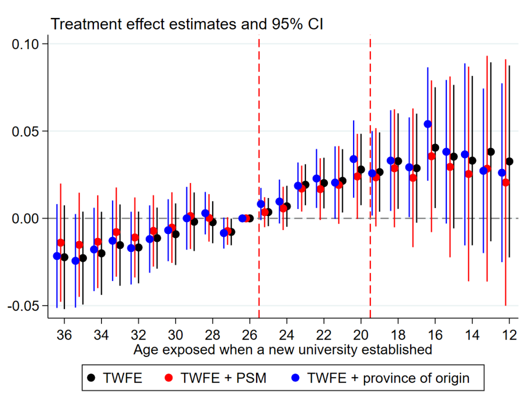
(c) Monthly wage



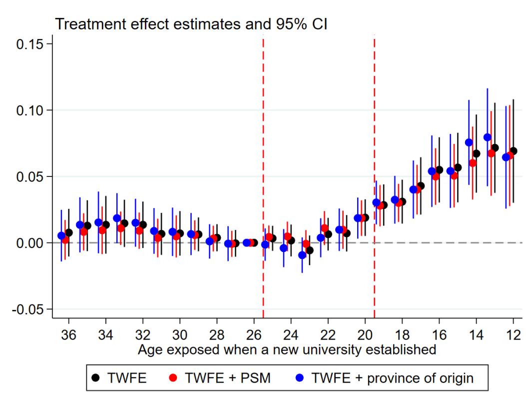
(d) Agricultural employment



(e) Manufacturing employment

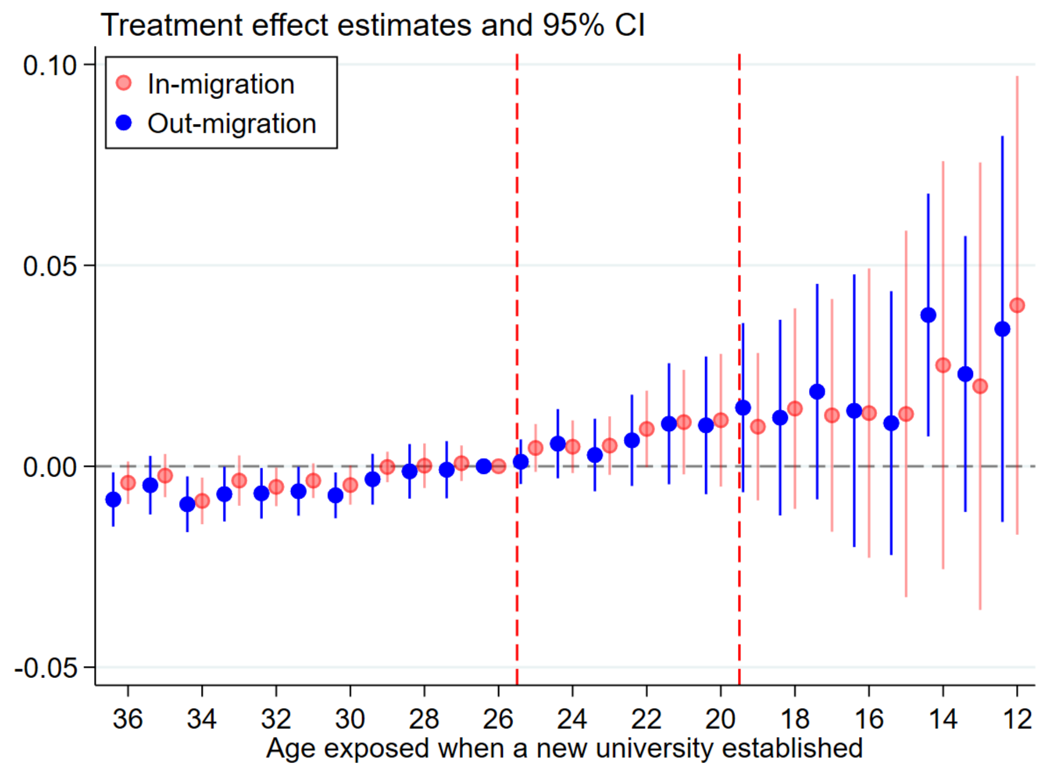


(f) Service employment



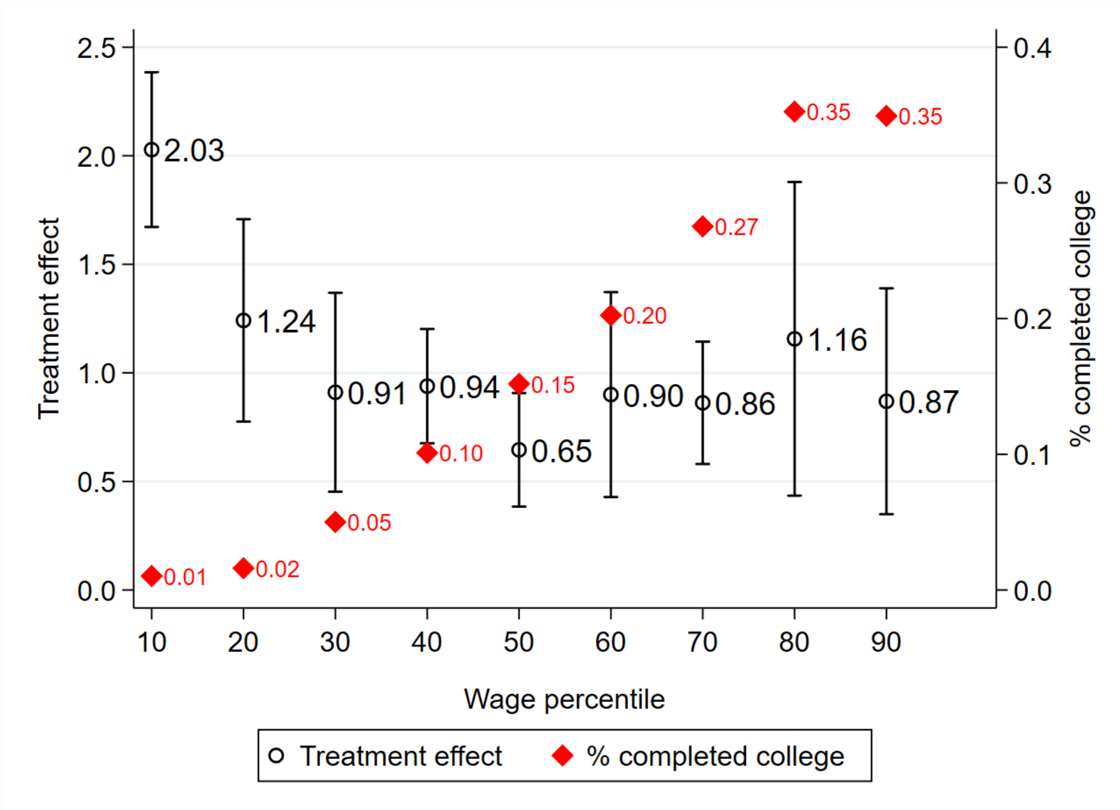
Note: Event study estimation for the effects of the higher education expansion on individual-level outcomes using different control groups. All models control for age, age squared, and gender. Standard errors are clustered at the province-level and 95% confidence intervals are displayed.

Figure A2: DiD estimate for the effects on in-migration and out-migration across provinces



Note: Graph shows the event study estimation for the effect of expansion on in-migration and out-migration. In-migration effect is measured by effect on migrant-status on a sample using province of current residence. Out-migration effect is measured by effect on migrant-status on a sample using province of origin.

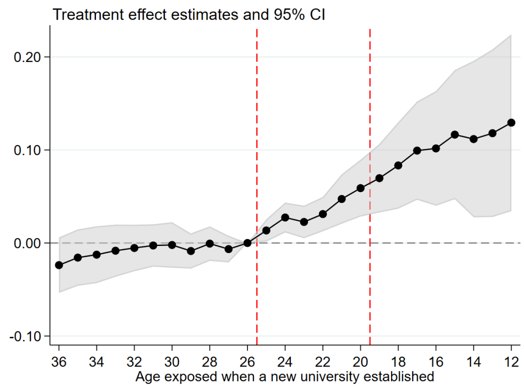
Figure A3: Change-in-changes Estimates for the effect of the expansion on log monthly wage



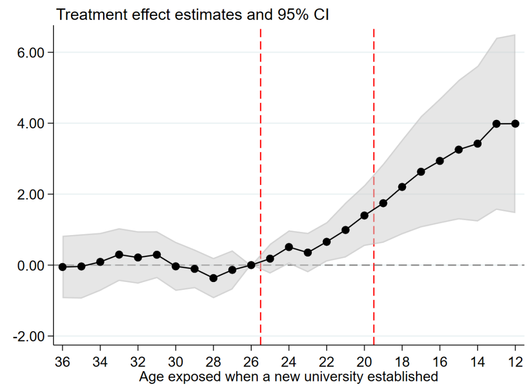
Note: The graph shows the distributional treatment effects estimated from a change-in-changes model (Athey and Imbens, 2006).

Figure A4: Event study estimation (cohort) for other labor market outcomes

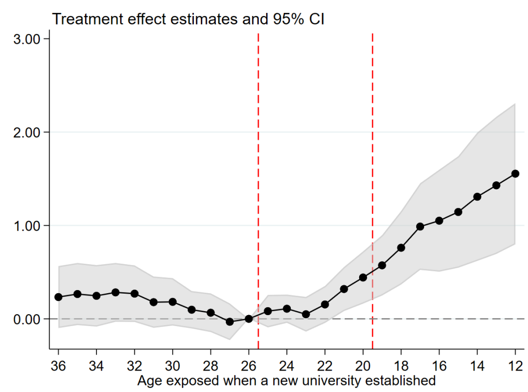
(a) Formal employment



(b) Occupational wage rank

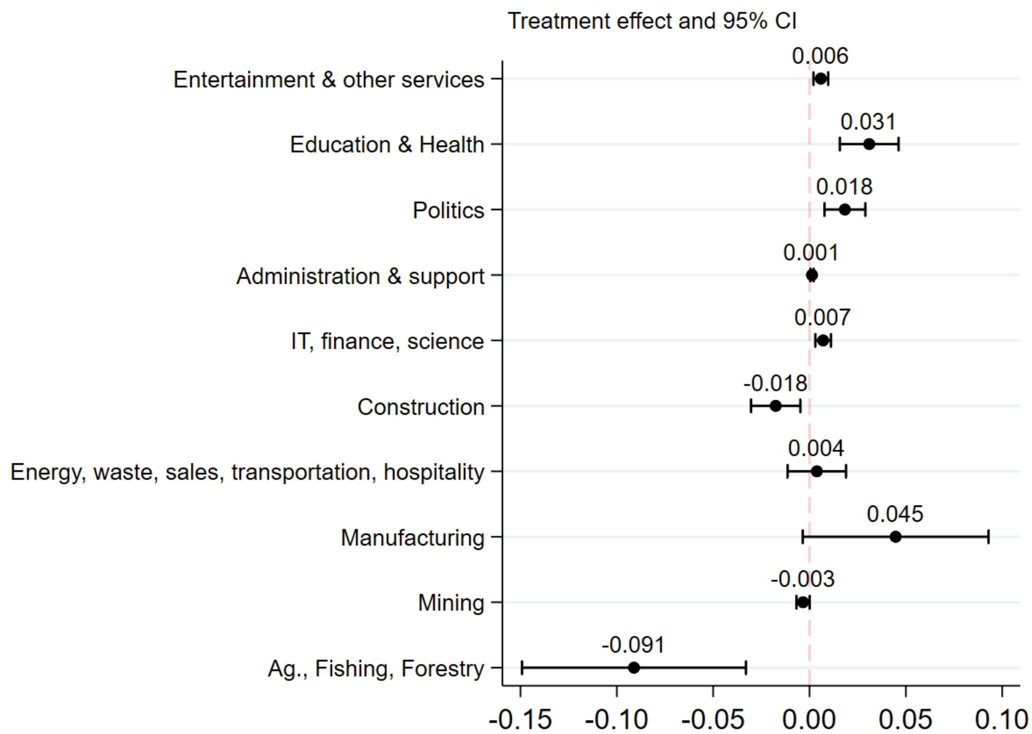


(c) Industry skill intensity rank



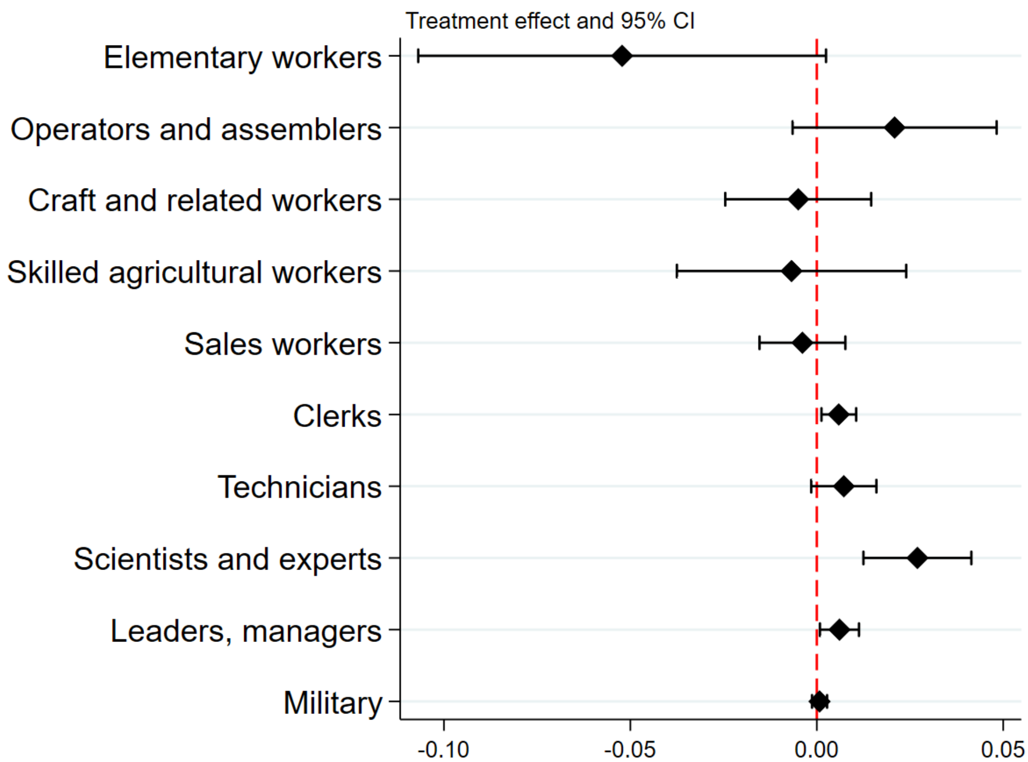
Note: The graphs display event study estimation for the effects of the higher education expansion on other labor market outcomes. Formal employment is defined as having a job that provides social insurance. Occupational wage rank measures how well-paid a given occupation is on the wage ladder; higher rank means having a higher-paid occupation. Industry skill intensity rank measures the how many college-educated workers in a given industry relative to others. All models control for age, age squared, and gender. Standard errors are clustered at the province-level and 95% confidence intervals are displayed.

Figure A5: DiD estimate for the effects on specific industry



Note: Graph shows the coefficient of the DiD interaction term and its 95% confidence interval. Each outcome is a binary variable indicating whether an individual works in the given industry. Standard errors are clustered at the province level.

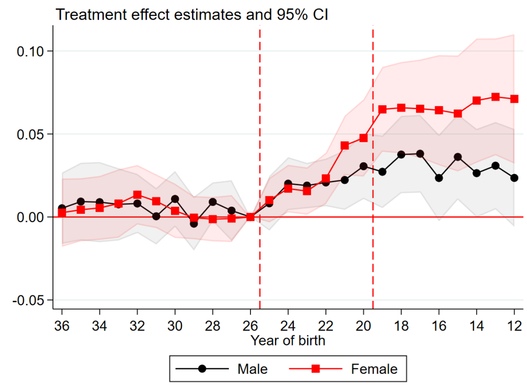
Figure A6: DiD estimate for the effects on type of occupations



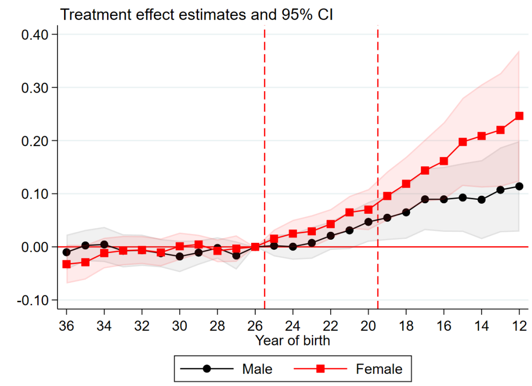
Note: Graph shows the coefficient of the DiD interaction term and its 95% confidence interval. Each outcome is a binary variable indicating whether an individual works in the given occupation. All models control for province-by-year fixed effects and birth cohort fixed effects. Standard errors are clustered at the province level.

Figure A7: Event study estimation (cohort) for the effects of the higher education expansion by gender

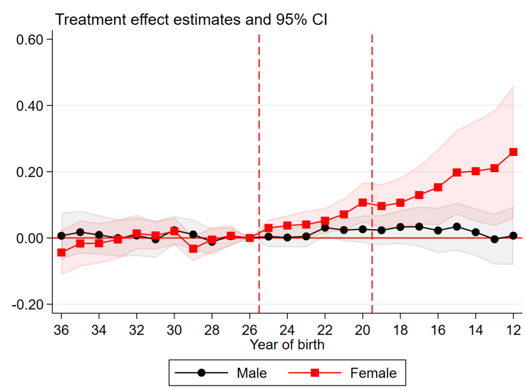
(a) College education



(b) Employment

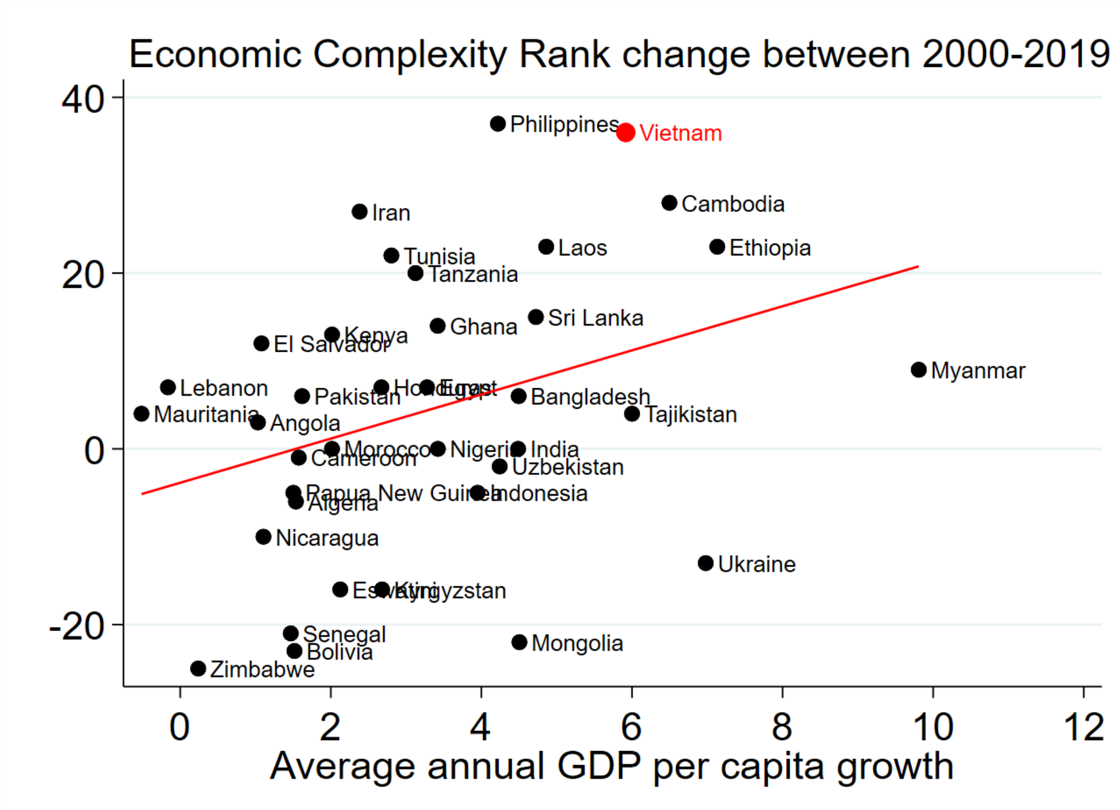


(c) Monthly wage



Note: The graphs display event study estimation for the effects of the higher education expansion on individual-level outcomes by gender. All models control for age and age squared. Standard errors are clustered at the province-level and 95% confidence intervals are displayed.

Figure A8: Economic Complexity Rank change and Annual GDP growth of lower middle income countries



Note: Annual growth in GDP per capita is from the World Development Indicators database provided by the World Bank. Economic complexity rank is publicly available in The Observatory of Economic Complexity database (Hausmann et al., 2014).

Table A1: Summary statistics by year and district

	Control		Treatment	
	2006-2011	2012-2018	2006-2011	2012-2018
Log TFP				
	2.512	2.530	2.530	2.541
	(0.141)	(0.141)	(0.131)	(0.143)
Labor productivity				
	17.686	17.962	17.812	17.986
	(0.972)	(1.133)	(0.943)	(1.231)
Capital-labor ratio				
	19.839	20.014	20.007	20.197
	(1.225)	(1.332)	(1.168)	(1.265)
Log total employment				
	2.431	2.364	2.330	2.222
	(1.322)	(1.530)	(1.235)	(1.464)

Table A2: Summary statistics by year and district

	Control		Treatment	
	2011	2015-2019	2011	2015-2019
% of adults who complete college or higher				
	0.064	0.086	0.157	0.206
	(0.054)	(0.073)	(0.110)	(0.127)
% of adults who are employed				
	0.325	0.400	0.526	0.596
	(0.136)	(0.165)	(0.125)	(0.110)
% of non-college adults who are employed				
	0.285	0.353	0.454	0.513
	(0.129)	(0.164)	(0.106)	(0.104)
% of college-educated adults who are employed				
	0.924	0.891	0.903	0.893
	(0.095)	(0.130)	(0.076)	(0.058)
Log monthly wage				
	10.111	10.299	10.329	10.643
	(0.218)	(0.362)	(0.227)	(0.257)
Log monthly wage of non-college workers				
	10.034	10.245	10.199	10.568
	(0.223)	(0.386)	(0.200)	(0.238)
Log monthly wage of college-educated workers				
	10.399	10.664	10.577	10.848
	(0.226)	(0.254)	(0.235)	(0.259)
Skill premium				
	0.360	0.408	0.378	0.280
	(0.252)	(0.363)	(0.193)	(0.179)

A Estimating total factor productivity at the firm level

Consider the following Cobb-Douglas production function for firm i in year t:

$$VA_{it} = \beta_l L_{it} + \beta_k K_{it} + \omega_{it} + u_{it}$$

where VA_{it} is the annual value-added, L_{it} is total labor, K_{it} is capital, measured as the value of assets at the beginning of the year (Newman et al., 2015), and ω_{it} is the unobserved productivity shock. Given that OLS is typically biased as both L_{it} and K_{it} are likely affected by the unobserved productivity shock, we first assume that firms' investment decision is a function of labor, capital, and productivity shock, i.e., $I_{it} = f_t(L_{it}, K_{it}, \omega_{it})$,[19] which makes ω_{it} observable in the production function (by inverting f_t):

$$VA_{it} = \beta_l L_{it} + \beta_k K_{it} + f_t^{-1}(L_{it}, K_{it}, I_{it}) + u_{it}$$

This approach by Ackerberg et al. (2015) (ACF) is different from two other convention approaches to estimate production functions, namely Olley and Pakes (1996) (OP) and Levinsohn and Petrin (2003) (LP), who do not include labor input as part of firms' investment decision. Not allowing labor input to enter the investment function means that L_{it} is a deterministic function of capital and investment and, hence, would be functionally dependent on the inverse function of investment; in other words, the coefficient of labor would not be identified.

Assuming that the productivity shock follows a first order Markov process, we can write $\omega_{it-1} = g(\omega_{it-1}) + \zeta_{it}$ where $g(\omega_{it-1})$ is the predictable component and ζ_{it} is the unpredictable/innovation component of productivity (Olley and Pakes, 1996). We also assume the following capital formation process: $K_{it} = (1-\delta)K_{it-1} + I_{it-1}$. These assumptions give us

[19]Investment is measured as annual change in value of fixed and long-term assets plus accumulated depreciation (Newman et al., 2015).

$E[\zeta_{it}|I_{it-1}] = 0$ and $E[\zeta_{it}|K_{it}] = 0$ (since K_{it} is determined at $t-1$). Lastly, we assume that $E[\zeta_{it}|L_{it-1}] = 0$ (Ackerberg et al., 2015). Given this set of moment conditions, we can estimate β_l and β_k.

Given that this approach requires panel data, we aggregate the firm-level variables to the district-by-year-by-industry level. We then estimate the production function by sectors and present the estimation results in Table A3. Once we obtain the estimates for β_l and β_k, we can use these estimates to calculate ω_{it} for each firm in each year, which is also our measure of total factor productivity (TFP).

Table A3: Production function estimation results

	Capital		Labor	
Agriculture				
	0.271***	(0.098)	1.143***	(0.392)
Mining				
	0.222***	(0.031)	1.063***	(0.048)
Manufacturing				
	0.342***	(0.031)	0.919***	(0.041)
Waste and electricity				
	0.380***	(0.065)	0.980***	(0.126)
Construction				
	0.198***	(0.028)	0.948***	(0.028)
Wholesale and retail				
	0.201***	(0.036)	1.386***	(0.055)
Transportation				
	0.057	(0.038)	1.103***	(0.055)
Hospitality				
	0.222***	(0.071)	0.682***	(0.186)
Information and communication				
	0.314***	(0.062)	1.035***	(0.105)
Finance, banking, and real estate				
	0.175***	(0.024)	1.348***	(0.088)
Science and technology				
	0.123***	(0.035)	1.384***	(0.050)
Administrative and support				
	0.156***	(0.041)	1.144***	(0.046)
Education, health, and social support				
	-0.116	(0.191)	1.535***	(0.296)
Entertainment				
	-0.112	(0.166)	1.677***	(0.268)
Other services				
	0.236	(0.250)	1.220**	(0.584)