Policy Research Working Paper 10193

Caring for Children and Firms?

The Impact of Preschool Expansion on Firm Productivity

Massimiliano Cali
Hillary C. Johnson
Elizaveta Perova
Nabil R Ryandiansyah

WORLD BANK GROUP
East Asia and the Pacific Region &
Macroeconomics, Trade and Investment Global Practice
September 2022

Abstract

Childcare services enable women who were previously unable to work due to taking care of their children to join the labor market. If some women are more productive in market work, rather than unpaid household work, the availability of childcare can potentially improve the allocation of talent across different occupations, triggering an increase in productivity. This paper tests this hypothesis using a survey of manufacturing plants and data on preschool expansion in Indonesia. The analysis relies on a triple difference estimation comparing plants in sectors with different degrees of female labor at baseline. The results suggest that between 2002 and 2014, when a rapid preschool expansion took place in Indonesia, an additional preschool per 1,000 children increased the total factor productivity of manufacturing plants by 11 percent for plants with an average fraction of female workers. The paper provides suggestive evidence that these effects were driven by better labor market matching, enabled by the expansion of female labor supply, and greater job stability for female employees. The results unveil a novel short-term economic impact of childcare services, which complements the long-term growth impact through human capital accumulation.

This paper is a product of the Office of the Chief Economist, East Asia and the Pacific Region and the Macroeconomics, Trade and Investment Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at eperova@worldbank.org.

Caring for Children and Firms? The Impact of Preschool Expansion on Firm Productivity[1]

Massimiliano Cali, Hillary C. Johnson, Elizaveta Perova, Nabil R Ryandiansyah

Keywords: childcare, productivity, misallocation, gender, Indonesia

JEL codes: J13, J16, D24

[1] We thank David McKenzie, Bob Rijkers, Achim Schmillen, Aart Kraay, Emcet Tas and participants of EAPCEO seminar for helpful discussions and suggestions. All errors remain our own.
Funding acknowledgement: The East Asia and Pacific Gender Innovation Lab is supported by UFGE in partnership with the Australian Department of Foreign Affairs and Trade. UFGE has received generous contributions from Australia, Canada, Denmark, Finland, Germany, Iceland, Ireland, the Netherlands, Norway, Spain, Sweden, Switzerland, the United Kingdom, the United States, the Bill and Melinda Gates Foundation, and the Wellspring Philanthropic Fund. Additional financial support for this paper was provided by the Government of Australia's Department of Foreign Affairs and Trade through the Australia-World Bank Indonesia Partnership (ABIP).

1. Introduction

The availability of quality early childhood education and development (ECED) programs yields substantive economic benefits in the long run by raising the human capital levels of future generations of workers (Garcia et al., 2020; Gertler et al., 2014; Auger et al., 2014; Herbst, 2017; Cascio, forthcoming). However, governments' investments, including in ECED, are typically driven by their shorter-term impacts. The key short-term economic benefit of ECED is the improvement in maternal labor market outcomes, including employment and earnings, as these programs also serve as a form of childcare.[2] While the evidence suggests that increased female labor force participation is associated with higher GDP (Bloom et al. 2009), it remains unclear whether the short-term benefits from public financing of childcare outweigh the costs. Barros et al. (2011) find that increases in household income due to access to childcare in Brazil are significantly lower than the cost of the program, which calls into question whether childcare is an efficient social protection instrument. In a similar vein, Andersen and Havnes (2018) find that gains in household income through increased maternal employment due to childcare availability in Norway did not increase tax income payments sufficiently to compensate the Norwegian state for the costs of childcare services.

Our paper contributes to this debate by testing for a different but potentially important short-term economic impact of childcare availability—the improved productivity of firms resulting from greater female labor supply. Although this economic externality of childcare has not been studied before, we argue that it is both theoretically plausible and critical to inform the decision on public financing of childcare. From a theoretical point of view, if ability and preferences over sectors, including the home production sector, are heterogeneously distributed, barriers to women's labor market choices can lower growth by creating a misallocation of talent and lowering the average talent pool of workers in the labor market (Hsieh et al. 2019). In a labor market framework with employer-employee matching and frictional

[2] See Morrissey (2017); Hegewisch and Gornick (2011); Del Boca (2015) for reviews of evidence from upper income economies and Halim, Perova, and Reynolds (2021); Harper et al. (2017) for reviews of evidence from developing countries. Olivetti and Petrongolo (2017) complement a review of microeconomic literature from high-income countries with cross-country analysis of spending on childcare (as percentage of GDP) and women's labor market outcomes. They conclude that both cross-country and micro data suggest that early childhood spending (as percentage of GDP) increases female employment.

search, relaxing these barriers will enrich a pool of potential matches for the firms, enabling them to match faster (Pissarides, 2000). Under certain conditions, this faster (and better) match may lead to increasing aggregate productivity. For example, Beerly et al. (2021) show that the increase in the available pool of workers due to easing of migration rules in Switzerland enabled firms to find workers who are better matches for available jobs and increased employment and wages.

We examine the link between ECED expansion and firm productivity in the context of Indonesia. This is a suitable laboratory for our analysis for several reasons. First, the female labor force participation (FLFP) rate in Indonesia is relatively low: 49% in 2002, the beginning of the period we study, compared to 70.6% in the rest of developing East Asia and Pacific (World Development Indicators, 2021). The FLFP rate remained low over our study period: by 2014, the FLFP rate in Indonesia only increased to 52.8% while the number declined to 66.47% in the rest of developing East Asia and Pacific. However, average rates conceal heterogeneous changes across the educational distribution. More educated women significantly increased their labor force participation, while less educated women at the lower end of income distribution withdrew from the labor market, as incomes grew (Schaner and Das, 2016). Second, starting in 2003, the Government of Indonesia adopted several pieces of legislation and increased public spending to facilitate public and private provision of ECED programs including preschools, which simultaneously provide early learning for children and serve as childcare. As a result, the gross pre-primary (age 3-6) enrollment rate more than doubled, jumping from less than 25 percent in 2004 to 60.3 percent by 2016 (World Development Indicators, 2021). Importantly, this growth has been geographically heterogeneous due to differential implementation across districts.

Such a dramatic increase in access to preschools enabled Indonesian women, who have relatively high levels of human capital, to join the labor market. Halim et al. (2022) show that an additional public preschool per 1,000 eligible children increased women's employment probability by 9.1 percent. This increase in labor supply may be particularly salient for Indonesian firms, which are more likely to cite inadequate skills as a barrier for hiring skilled workers compared to other neighboring countries, including Malaysia and Thailand. According to World Bank Enterprise Survey data, 55% of Indonesian firms report this constraint compared to 39.5% and 46% of Malaysian and Thai firms, respectively.

To perform the analysis, we combine a large annual panel of manufacturing plants (Statistik Industri – SI henceforth) for the period from 2002 to 2014 with the village census data (PODES), which contains information on the number of preschools available per village. As village identifiers are not available in the Manufacturing Survey, we aggregate the number of preschools at the district level, taking advantage of the representativeness of firm-level data at such low level of geographic disaggregation. We use a triple-difference framework exploiting variation in preschool expansion geographically and across time, as well as sector-specific baseline reliance on female labor supply to identify the impact of preschool availability on plants. We argue that plants that relied more on female labor at baseline would benefit more from preschool expansion. This relies on the assumption – which we confirm empirically - that preschool expansion is more likely to affect female than male labor supply given women's greater role in childcare responsibilities. This approach allows us to control for all time-varying confounders at the district and sector levels, thus addressing the main concerns to the causal interpretation of the estimates.

We find a significant positive impact of the expansion of preschools on plant-level total factor productivity – whether measured in revenue (TFPR) or quantity (TFPQ) terms. Our preferred specification suggests economically substantial impacts: an increase of 1 additional preschool per 1,000 eligible age children (approximately 1/3 of standard deviation) increases TFPR by 11 percent for plants in sectors with 41 percent female employees (which is the average in our sample). We also explore two mechanisms through which institutional childcare may impact firm productivity via the increased female labor supply: (i) improvements in allocation of talent across jobs; and (ii) increase in job stability due to reduction in turnover.

First, preschools may alleviate one of the constraints that women face when joining the labor market - childcare responsibilities - enabling them to choose between productive and reproductive work based their ability and earning potential. This is likely to improve allocative efficiency of talent across jobs, potentially leading to an increase in productivity. We provide evidence in support of this mechanism by demonstrating that preschool expansion had a positive impact on the share of women hired by manufacturing firms using SI data. We complement this evidence with an analysis based on individual Labor Force Survey data (SAKERNAS), which shows

that preschool availability increases the probability of women to be employed, particularly those with small children.

The second mechanism that we explore is related to the evidence that availability of childcare services increases job stability (Angeles et al., 2011) and employee attendance (Ranganathan and Pedulla, 2020). That is because it provides more flexibility in terms of working arrangements to workers with childcare responsibilities. For the plants, lower turnover in turn would raise productivity (Brown, 1989; Brown and Medoff, 1978; Shaw, 2011; Strober, 1990; Yanadori and Kato, 2007; Giesing and Laurentsyeva, 2017; Kacmar et al., 2006; Grinza and Rycx, 2020). We find suggestive evidence in support of this channel using SAKERNAS data: expansion of preschool availability reduces the turnover of women in manufacturing, particularly for women living in households with small children.

A further implication of both channels is that wages may increase as a result of increased labor supply. Unlike in a classical supply and demand framework in a perfectly competitive labor market, access to a larger pool of workers may enable the firms to find better matches between workers and jobs, possibly leading to increase in wages (Pissarides, 2000; Beerly et al. 2021). Similarly, increased job stability may lengthen the average tenure of workers in the firm, which would also increase average wages. We test these implications using SI data and find that indeed an increase in preschool density had positive impact on the average worker's wages.

Our paper contributes to several strands of literature. First, we add to the literature on allocation of talent and productivity, building on Hsieh et al. (2019), by empirically demonstrating the role of one specific barrier to allocative efficiency between the market and home production sector: women's disproportionate childcare responsibilities. Second, we contribute to the empirical evidence testing Pissarides's theory of imperfect labor market with firm-employee matching and frictional search. Beerly et al. (2021) recently demonstrated empirical evidence to this theory in the context of lowering barriers to migration. We expand it to a new context: lowering barriers due to the introduction of childcare services, and test for the possibility of an increase in productivity (which Beerly et al. (2021) hypothesize takes place but do not show empirically). Third, we contribute to the literature on public investment in childcare. While investment in childcare has been argued for as cost-effective based on long-term gains in human

capital, our work offers an additional argument for it, demonstrating short-term impacts on firm productivity. Considering that in the aftermath of the COVID-19 pandemic many countries are grappling with a dual problem of mitigating human capital losses and boosting short-term productivity, our paper has a strong potential to inform policy debates.

The paper is organized as follows. Section 2 describes preschool expansion in Indonesia. Section 3 discusses the data and outlines the empirical strategy. Section 4 presents the results. Section 5 explores the channels behind observed impacts. Section 6 concludes.

2. Expansion of preschools in Indonesia

In the 1990s and 2000s, Indonesia was lagging the world average as well as its regional neighbors in terms of preschool enrollment (Figure 1). Indeed, in 2002 Indonesia's gross preschool enrollment of 24.4 percent amounted to about 64 percent and 72 percent of EAP and world averages, respectively. It was 31 percent lower than in the Philippines, 45 percent lower than in Vietnam, and 53 percent lower than in Malaysia (World Development Indicators, 2021). To improve conditions for its children and relying on international evidence on the role of early education for human capital accumulation, the Government of Indonesia has implemented a set of policies and programs that prioritize ECED.

Specifically, in 2003, Indonesia adopted ECED into its national education system. This measure provided the legal foundation for ECED to be considered in the national and regional education budget allocations. In 2004 inclusion of ECED into the Ministry of Education and Culture's Strategic Plan (Rencana Strategis or RENSTRA) and the Ministry of National Development Planning of the Republic of Indonesia (BAPPENAS) 5-year development plan provided further impetus for expansion of preschools. Block-grants were set up to incentivize private provision; district-level governments were responsible for the public provision of preschools. In 2005, the government issued a policy stipulating a minimum allocation of 20 percent of the annual national and regional budget to education expenses.

These reforms affected all types of ECED establishments; however, we focus on preschools due to data availability.[3] Preschools are non-mandatory, formal ECED establishments intended for children between the ages of 4 and 6. They offer academic preparation for primary education. Figure 2, Panel B shows the upward expansion of preschools overtime. Between 2003, when the first reforms were enacted, and 2014, the average number of preschools per 1,000 children of eligible age almost doubled, from 3.7 in 2003 to 6.6 preschools in 2014. As a result, the preschool enrollment rate (3-6) years more than doubled during the same time, from 25.7 percent in 2003 to 57.9 percent in 2014 (World Development Indicators, 2022). Figure 2, Panel A illustrates the geographical variation of preschool access across the country. In the next section, we elaborate how we exploit spatial and temporal variations in preschool access to estimate its impact on firm productivity.

3. Empirical strategy and data

3.1. Data

We combine several data sets to estimate the impact of access to preschools on firm-level outcomes. We use PODES, the village census, and the annual National Socioeconomic Survey (SUSENAS) to construct a measure of preschool density: the number of preschools per 1,000 children of eligible age (3 to 6 years old). Although preschools are intended for children aged 4 to 6, following Halim et al. (2022), we define preschool density with a broader age range, as age requirements are not strictly enforced. PODES is fielded approximately every three years. We use data from 2002 to 2014. SUSENAS is annual, and we merge corresponding years to PODES data. We measure preschool availability at district, rather than village level, as this is the lowest level geographical identifier available in all datasets that we use. Since decentralization reform in 1999, districts have often split over time. In 1993, there were 290 districts; by 2014, there were 511 districts. To ensure equal comparisons across time, we harmonize district boundaries as they existed in 1993. As PODES is not available annually, we follow Halim et al. (2022) and infer preschool data from in-between PODES years using the closest upper year available. For instance, year 2007 is sandwiched between PODES 2005 and 2008, so we infer preschool data from the

[3] Specifically, we are looking at *Taman Kanak-kanak* (TK) and *Raudhatul Afthal* (RA).

2008 round. We check robustness of results to two alternative strategies to account for the missing years.

We use plant level outcomes from the Indonesian survey of manufacturing plants with at least 20 employees Statistik Industri (SI henceforth). SI is administered by Indonesia's Badan Pusat Statistik (Central Agency on Statistics). The survey has extensive coverage: it is administered as a census every 10 years starting from 1996 and is very close to a census in the remaining years. Such coverage ensures high representation even at low levels of sectoral and geographic disaggregation. The data allow for different groupings of firms into sectors, based on definitions of Klasifikasi Baku Lapangan Usaha Indonesia (KBLI), a classification mostly compatible with ISIC Rev.3. We use 5-digit sector levels as our primary specification; however, our results are also robust to an alternative specification using 3-digit sector classifications. For the main analysis we use data from 2002 to 2014, the last year of data which allows us to compute TFPQ estimates. We also use the change between 2000 and 2002 to check our identifying assumptions.

Following Cali et al. (2021) we use SI data to estimate the plant-level performance measures: TFPR, TFPQ and markup. Unlike most existing studies using industry-level price indexes to deflate nominal variables, the granularity of our data allows us to calculate plant-specific output and input price indexes. This mitigates the bias deriving from inputs' price heterogeneity across plants and allows us to disentangle technical efficiency (TFPQ) and markups. To estimate plant-level production function parameters, we follow the control function approach and timing assumptions of Ackerberg et al. (2015), which are useful to mitigate the simultaneity bias affecting simple OLS coefficients. We slightly modify the standard estimator to allow for the presence of adjustment costs in materials. Then, using the estimated production functions, we calculate plant-level markups based on a model of profit maximization by firms, as in De Loecker and Warzynski (2012). Finally, we combine our estimates of TFPQ and markup to obtain a measure of TFPR. Appendix I elaborates construction of these measures.

SI also provides additional time-variant control variables: raw materials and share of foreign ownership. We use variation in exposure to changes in preschool density to identify impacts of childcare on firm productivity. We construct a measure of exposure based on SI data using the average fraction of female employees in the firms in the same sector in 2002 - prior to

preschool expansion. Measures of worker income and the number of workers enable us to test some of the mechanisms behind the identified impacts.

We also explore mechanisms utilizing Indonesia's national labor force survey (SAKERNAS). SAKERNAS is a nationally representative survey of all working-age individuals conducted bi-annually in February and August. We use the August rounds of SAKERNAS data (for the 2007-14 period), which are representative at the district level, to construct indicators for employment in manufacturing, services, and wage employment for men and women, as well as an indicator for a change in employment during the 12 months prior to data collection.

Lastly, we employ data from the Indonesia Database for Policy and Economic Research (INDO-DAPOER) to conduct robustness checks. This data set was constructed by the World Bank from several governmental data sets and provides us with district-level economic and social indicators. We use total government per capita expenditure as a measure of fiscal development. To proxy socio-demographic dynamics, we take the district-level average scores on national exams of lower secondary and upper secondary students. Finally, we use the length of highways as a fraction of total road length to proxy quality of infrastructure. These variables enable us to test whether preschool expansion is correlated with other economic or social changes that may drive the observed impacts in firm productivity.

Table 1 provides summary statistics for outcome, treatment, and control variables from all the data sets.

3.2. Empirical strategy

Our identification of the impact of preschool availability on plant-level outcomes relates changes in performance of plants with varying degrees of exposure to changes in preschool availability. Specifically, we compare plants along two dimensions. The first is the degree of reliance on female workers. To that end we distinguish sectors (at a granular level of classification) with higher vs. lower share of female workers in the same district – hence exposed to the same changes in preschool density over time. The former should benefit more from any given level of preschool expansion, as we expect the absence of childcare services to primarily

constrain women's (and not men's) labor. Computing the female shares at the sector- rather than at the plant-level allows to reduce the endogeneity concerns of this exposure measure.

The second dimension is the exposure to different degrees of preschool expansion: as Figure 2 demonstrates, there is a substantial variation in access to preschools across time and districts. For any given initial share of female workers, the plants located in districts with a greater increase in preschool expansion are more likely to increase their performance due to the related labor supply shock.

More formally, we estimate a triple difference regression of the type:

$$Y_{idst} = \beta FemShare_s^{t_0} \times Preschool_{dt} + \mathbf{\Gamma X_{idt}} + \alpha_i + \gamma_{dt} + \delta_{st} + \varepsilon_{idst} \quad (1)$$

where Y_{idst} is the outcome variable of manufacturing plant *i* in district *d* and sector *s* in year *t.* Our key outcome variables include TFPR and TFPQ, which measure plant-level productivity in revenue and quantity terms, respectively, and markup. TFPQ captures the pure technical efficiency of plants as it carves markup out of the TFPR (see Appendix I). α_i, γ_{dt}, and δ_{st} are plant fixed effect, district-time and sector-time fixed effects, respectively. $\mathbf{X_{idst}}$ is a vector of time-varying plant level controls (raw materials and share of foreign ownership). The variable $Preschool_{dt}$ captures preschool access in district *d* at time *t,* measured as the number of preschools per 1,000 children of eligible age, as described above. We expect preschool availability to affect plants differently, depending on how strongly they rely on female labor. Thus, we interact preschool density with $FemShare_s^{t_0}$ - the share of female workers in sector *s* at t_0 (2002 or in the first year the plant entered the manufacturing survey, whatever comes later).[4]

The inclusion of plant, district-time, and sector-time fixed effects addresses two primary threats to the causal identification of the effects of preschool availability on plant productivity. First, preschool expansion may be correlated with other governmental programs which may affect firms' performance. For example, if the Ministry of Education targets a subset of districts for all its programs, from preschool to tertiary, preschool access may be correlated with the

[4] To minimize possible remaining endogeneity concerns, we exclude the plant itself from the computation of the sector-specific shares applied to each plant.

presence of vocational programs or the availability of higher human capital due to stronger schools and tertiary institutions.[5] District-time fixed effects should address this concern.

Second, the sector-specific proportion of women among employees may be correlated with other characteristics that matter for performance, both across plants and across sectors over time. For example, nearly everywhere in the world women tend to work in lower productivity sectors. Similarly, the evolution of performance over time may be different for female-dominated sectors, and this could also affect the demand for preschools. The inclusion of plant and sector-time fixed effects should absorb this unobserved heterogeneity across plants, sectors and time.

Thus, once we control for sector-time, district-time, and plant fixed effects, it is plausible that the observed differences in productivity associated with an additional preschool per 1,000 children between plants with higher and lower shares of female workers are due to the impact of preschools: plants with a higher share of female workers benefit from a higher "dose" of preschool treatment.

However, the possibility of endogenous placement of preschools remains a concern. For example, the government may incentivize preschool construction in districts with high levels of female-dominated industries with increasing productivity. Alternatively, the private sector may respond to demand for childcare in districts with large share of highly productive firms in female concentrated industries – as they expect women working in such firms to have higher demand for childcare, and be able to cover its costs. To address this concern, we carry out two separate checks for parallel trends. First, we carry out a placebo test using the data from 2000, 2001 and 2002, prior to extensive preschool expansion, and estimate:

$$Y_{idst} = \beta FemShare_s^{t_0} \times Growth_d \times Post_t^p + \mathbf{\Gamma X_{idt}} + \alpha_i + \gamma_{dt} + \delta_{st} + \varepsilon_{idst} \qquad (2)$$

where $Growth_d$ captures change in preschool density between 2003 (when the NSEA was passed) and 2014. $Post_t^p$ is a dummy equal to 1, if year *t* is greater or equal to year *p.* We estimate

[5] It could also be the case that the election of a particularly efficient district-level government would benefit firms, for example due to a better local business climate, and attract private preschool investments at the same time.

equation (2) for two values of *p*: 2001 and 2002. The remainder of notation is the same as in equation (1).

Table 2 presents the results. We do not find evidence of differential change in productivity across plants from sectors with different shares of women in response to changes in preschool density when using 2001 as a cut-off point (columns 1-3). When we use 2002 as a cut-off (columns 4-6), the coefficients on TFPR and TFPQ are significant but negative: suggesting that preschools could be expanding in districts where plants in female-dominated sectors were experiencing declining productivity. Hence if we find any impact in the opposite direction, it will likely be an underestimate.

Our second check focuses on the relationship between changes in plant productivity prior to preschool expansion and changes in preschool density across plants from sectors with different shares of female employees. Specifically, we estimate:

$$\Delta Y_{ids} = \beta FemShare_s^{t_0} \times \Delta Preschool_d + \mu_d + \theta_s + \varepsilon_{idt} \quad (3)$$

where $\Delta Y_{ids} = Y_{ids}^{2002} - Y_{ids}^{2000}$, change in productivity outcomes between 2000 (the first year available in our data), and 2002, the year before preschool expansion began. $\Delta Preschool_d = Preschool_d^{2014} - Preschool_d^{2002}$ captures changes in preschool density between the year prior to preschool expansion and the last year available in our data. μ_d and θ_s are district and sector fixed effects, respectively. Table 3 confirms that changes in plant productivity prior to preschool expansion are not positively correlated with preschool expansion, depending on the share of women in the sector. The coefficient is negative and not significant in case of TFPR, and negative and significant at 10% level for TFPQ. Again, preschools appear to expand in districts where firms with high fraction of female employees experienced decreases in productivity prior to their expansion. Capturing a positive impact would suggest that availability of preschools has reversed a negative trend.

4. Results

4.1. Main results

We begin by estimating the impacts of preschool expansion on plant performance, captured through the TFPR, TFPQ and markup variables. Table 4 shows positive and strongly significant impacts of preschool on productivity. For example, an increase of 1 additional preschool per 1,000 children of eligible age (approximately 1/3 of standard deviation) triggers a 11 percent increase in TFPR for plants in sectors with 41 percent of female employees (which is the average in our sample). For a plant from a sector which employs only women (0.4 percent of firms in our sample), the increase in TFPR would be 30 percent, and for a plant in a sector where only 9-11 percent of employees are women (3 percent of firms in our sample), productivity would go up by approximately 3 percent.[6] The impacts on TFPQ are very similar: building an additional preschool per 1,000 children triggers an 11 percent increase in TFPQ for a firm from a sector with average exposure, or 41 percent of women among employees.

These results are consistent with zero impact on markup: coefficients are low in magnitude and not statistically significant. Indeed, the TFPQ measure captures changes in pure technical efficiency of the firms, carving markup out of the TFPR measure. Combined, these coefficients suggest that access to preschools increases productivity through a boost in technical efficiency rather than product differentiation or market power.

4.2. Robustness

We check the robustness of the results to several remaining threats to our identification. First, the baseline results rely on a measure of a plant's dependence on female work at the sector-level rather than at the plant-level. While this choice minimizes endogeneity concerns, it also throws away part of the variation in female work dependence across plants. To address this issue, we re-estimate equation (1) by computing the share at the plant level ($FemShare_i^{t_0}$) and instrumenting it with the average share at the sector level $FemShare_s^{t_0}$, excluding plant *i*. Table 5 shows the results. The power of the instrument is very high with the reported F-statistics exceeding the recommended minimal threshold (Stock et al., 2002) by orders of magnitude (they

[6] There are no firms with 10% female employees.

range from 1,282 to 1,315, depending on the outcome of interest). The coefficients instrumented with the IV approach remain strongly significant and barely change in magnitude.

Second, we check if the results are robust to using a less disaggregated sectoral classification than 5-digit KBLI to compute the female share of workers at baseline. Table 6 shows that the results are robust to using the 3-digit KBLI sector specification, for both OLS and IV estimations.

Third, we also check the robustness of our results to using two alternative methods of calculating preschool density during the years that are missing in PODES. The first consists of restricting our constructed panel to PODES years. The second, is to predict preschool density for the missing years using linear projection with the closest two data points available. For example, to estimate the preschool density in 2007, we fit a linear projection using preschool density data in PODES 2005 and 2008. Table 7 presents the results. The coefficients on TFPR and TFPQ remain strongly significant. They are 5 and 8 percentage points lower in magnitude, respectively, when we use reduced sample (columns 1-3), and 6 and 2 percentage points higher in magnitude, when we use linear projection (columns 4-6). Notably, the coefficient on markup becomes significant at the 10% level when linear projection is used to impute missing years. However, overall, the results are similar in direction of effect and in magnitude.

Finally, we address a further concern to the robustness of our estimates, i.e. that some district-level time-variant variables, potentially correlated with preschool expansion, may affect plants differently depending on the share of female workers. For example, if investment in education or overall changes in district-level infrastructure, such as road improvement, affect plants differently depending on the share of women, our estimates in (1) may be biased, even though we control for district level time-variant heterogeneity with district-time fixed effect γ_{dt}. Consider the following scenario: preschool expansion may be correlated with changes in local public sector efficiency more generally. Such district-level change may differentially affect firms with different shares of female workers, if, for instance, higher ability managers employ more women and are more able to take advantage of greater efficiency of the local public sector.

We address this concern by including in equation (1) interactions of $FemShare_s^{t_0}$ with a set of district-level time-variant variables, which may plausibly affect the productivity of plants with smaller or larger shares of female employees differently. These district level controls aim to reflect changes in public sector funding, in overall infrastructure, and in education outcomes. Inclusion of interactions will purge such confounding effects from our coefficient of interest. Insignificant coefficients on interactions between $FemShare_s^{t_0}$ and district-level controls would signal that firms with a higher/lower share of female workers do not benefit differently from district level changes in public spending, infrastructure and education. Specifically, we estimate:

$$Y_{idst} = \beta FemShare_s^{t_0} \times Preschool_{dt} + \mathbf{TZ_{dt}} \times FemShare_s^{t_0} + \mathbf{\Gamma X_{idst}} + \alpha_i + \gamma_{dt} + \delta_{st} + \varepsilon_{idst} \qquad (4)$$

where vector $\mathbf{TZ_{dt}}$ includes three variables: total government expenditure per capita, a measure of road infrastructure and a measure of the quality of education in the district, as described in Section 3.1.

Table 8 presents the estimates from equation (4). Coefficients in columns (I), (II), (V) and (VI) suggest that there is no variation in the impact of government expenditure or road infrastructure on TFPR and TFPQ, depending on the share of female workers per firm. Notably, the coefficient on the interaction between female share and an indicator of education quality is statistically significant. Plants with a higher share of female workers are able to take greater advantage of educational attainment in their districts. This is not surprising: women out of labor force in Indonesia are generally higher educated than men who remain out of labor force. Thus, by employing more women firms may be able to tap into greater pool of highly qualified potential workers. Although growth in overall level of education may be positively correlated with preschool expansion, inclusion of this additional covariate barely changes the magnitude of our main coefficient of interest – the change is less than 0.1 percentage point.

5. Channels

We explore two possible mechanisms that can lead to the observed impacts of preschools on plant productivity. First, following Hsieh et al. (2019) we check whether the observed increase

in productivity may be due to better sorting across sectors, where home production is one of the sectors. Second, we explore the possibility that preschool availability may lower worker turnover among female employees, which also may result in productivity gains. Both channels are consistent with an increase in wages of the plants' workers. Hence as a further implication of both channels, we check also whether the increase in preschool availability leads to an increase in workers' wages, consistent with Pissarides (2000) and Beerli (2021).

5.1. Improvement in sorting across sectors

Hsieh et al. (2019) look at how improved allocation of talent increases aggregate productivity. Their paper introduces forces such as group-specific social norms to the standard Roy (1951) model which act as distortions to the allocation across occupations. If individuals sort not only based on their ability or preferences, as is the case in the standard Roy (1951) model, but also on the basis of other factors, such as discrimination or social norms about who should take care of the children, such sorting reduces the average quality of workers in an occupation, and hence productivity. Hsieh et al. (2019) demonstrate that between 20 percent and 40 percent of US growth from 1960 to 2010 was driven by improved allocation of talent across occupations, including women leaving the home sector for market work.

Lack of childcare may be one such barrier that lowers the likelihood that women sort into market work. Removing this barrier is likely to bring qualified women into the labor market, increasing productivity. We use two data sets to test how preschool expansion impacted female employment. First, we rely on the manufacturing survey to estimate the impact of preschool expansion on employment of women in the plants in our sample. Specifically, we estimate equation (1) with the number and share of female workers in a plant as a dependent variable. We use the sample of all manufacturing plants, and then limit it to the sample for which TFPR and TFPQ measures are available. Table 9 presents the results. One additional preschool per 1,000 children in a district increases the share of female workers by 0.4 percentage points, or 1 percent, at the plants with average share of female workers (41 percent). The coefficient is significant at the 1% level. The results are very similar for both the full sample and the sample restricted to plants for which TFP measures are available. We also find evidence of impact on the number of female workers. However, these results are weaker as the coefficients are estimated

with a large margin of error. Our preschool interaction is only significant at 15% (t-statistic is 1.61) in the larger sample (column 3). One additional preschool per 1,000 children appears to increase the number of female workers by 1 percent. The coefficient remains positive but becomes instead close to zero in the restricted TFP sample (column 4).

Second, we corroborate the finding that plants are more able to tap into female workers' skills by using the data from SAKERNAS. We use repeated cross-section, spanning the years from 2007 to 2014, and estimate individual-level regression:

$$y_{idt} = \theta Female_i + \varphi Female_i Preschool_{dt} + \mathbf{\Gamma X_{idt}} + \gamma_{dt} + \epsilon_{idt} \quad (5)$$

where y_{idt} is a labor market outcome for individual *i* in district *d* at time *t.* We analyze three binary outcomes: employment in manufacturing, employment in services and wage employment. $Female_i$ is a dummy variable equal to 1 if the respondent is a woman, and $Preschool_{dt}$ is defined based on PODES data: number of preschools per child in district *d* at time *t.* $\mathbf{X_{idt}}$ is a vector of time-variant variables, which include education, experience, and experience squared. γ_{dt} is a district-time fixed effect.

While we cannot establish a causal impact, we can explore associations between preschool availability and women's employment. Table 10 shows that being a woman is negatively associated with employment in manufacturing, services and wage employment. However, the availability of preschools in the district appears to reduce this gender penalty. The coefficient on interaction between $Female_i$ and our measure of preschool density is positive and strongly significant. An increase in one standard deviation (0.003 additional preschool per child) increases the likelihood of employment in manufacturing by 0.27 percentage points more for women.

We also limit the sample to women, and check whether preschool availability affects women's employment differently, depending on the presence of children under 10 in the household.[7] We estimate:

[7] SAKERNAS only establishes the presence of children in the household, not the relationship with female respondents. It also cannot be used to limit the children's age to preschool eligibility age: under 10 is the best available proxy afforded by our data.

$$y_{ihdt} = \mu YoungChild_h + \omega YoungChild_h Preschool_{dt} + \mathbf{\Gamma X_{idt}} + \gamma_{dt} + \epsilon_{idt} \qquad (6)$$

where $YoungChild_h$ is equal to 1 if there is a child under 10 in the household *h* of woman *i.* The remainder of the notation is the same as in equation (5). Similarly, we find a significant negative association between the presence of young children in the household and the likelihood that a woman works, which is mitigated by preschool availability (Table 10, Columns 4-6).

5.2. Reduction in worker turnover

In addition to enabling labor market entry, preschool availability may increase firm productivity by increasing job stability, reducing worker turnover among women. Although the majority of studies which evaluate the impact of childcare services on maternal labor market outcomes do not analyze turnover, those that do suggest that access to childcare strengthens maternal attachment to jobs. Angeles et al. (2011) find that *Programa de Estancias Infantiles* in Mexico, which provides access to childcare for working mothers, increased job stability among its beneficiaries. Ranganathan and Pedulla (2021) find that availability of employer-sponsored childcare increases daily work attendance, which may contribute to a reduction in turnover.

Several studies, in turn, suggest that higher turnover is likely to result in productivity losses: high turnover deteriorates firm-specific stock of knowledge and expertise directly through the loss of trained workers and through lowering incentives for the firms to invest in training of their employees. This effect is captured by TFP, as this form of human capital cannot be fully accounted for in wages (Brown and Medoff, 1978; Shaw, 2011; Strober, 1990; Yanadori and Kato, 2007; Giesing and Laurentsyeva, 2017).

We explore the likelihood of this mechanism by examining the association between preschool availability and worker turnover using SAKERNAS data. We estimate regressions (5) and (6) with dummies equal to 1 if worker *i* changed job in year *t*. We carry out similar analysis for the likelihood of job turnover in manufacturing and services. Table 11 shows that women experience higher job turnover in manufacturing but lower job turnover in services and overall. Notably, a relatively low fraction of Indonesians is employed in manufacturing: about 10 percent of the full SAKERNAS sample. However, the availability of preschools in the district is significantly associated with a reduction in women's probability of changing manufacturing jobs within a year.

We do not find a similar association for services or overall job turnover. When limiting the sample to women only, we find that women in households with children under 10 are more likely to experience job turnover when employed in manufacturing. This relationship is reversed for services and overall job turnover. Again, we find that preschool availability is associated with a lower likelihood of job turnover for women in households with young children overall, and in manufacturing, but not in services.

5.3. Does increased supply of female workers lower wages?

A classical supply and demand framework in a perfectly competitive labor market predicts that a positive labor supply shock due to childcare provision may trigger a reduction in wages. In contrast, an increase in wages would be consistent with both channels tested in the previous section. Pissarides (2000) shows that in an imperfect labor market with firm-employee matching, frictional search and job posting, a positive labor supply shock may enable firms to match workers to jobs more effectively. Such improved matching can increase wages. Similarly, lower turnover may lengthen the average tenure of workers in the firm, which would also increase average wages given the positive relation between tenure and wages (Brown, 1989).

In table 12 we use the SI to test the impact of preschool expansion on worker wages by estimating Equation (1) with worker income as the dependent variable. We indeed find strong impacts on wages: an additional preschool per 1,000 children raises worker wages by approximately 1.5 percent in a plant with an average fraction of female employees (Table 12). The estimates are significant at the 1% level and of the same magnitude regardless of whether we use the full sample of plants (column 1) or limit the sample to plants for which the measures of TFPR and TFPQ are available (column 2).

To the extent that preschool expansion relaxes firms' skill availability constraint, these impacts should be larger precisely for those plants that are initially most constrained. To test for this differential effect, we exploit a question in the 2006 SI data that asks if the lack of workers' skill is the primary reason not to expand. Column (3) shows that those plants which answered 'yes' to this question increased wages as a result of preschool expansion significantly more than

the others.[8] In fact this difference in impact is larger than the average marginal impact of preschool on wages. This differential impact holds also in the restricted 'TFP sample' although it is only significant at 15 percent (column 4).

Taken at their face values these tests provide suggestive evidence in support of preschools raising manufacturing plants' productivity by enabling an expansion of the local female labor supply. This in turn allows better employer-employee matching and a reduction in turnover for female workers, which benefit particularly manufacturing plants in sectors relying more on female labor. As a result, the plants' share of female workers increases and so do average wages, especially in skill constrained plants at the beginning of the period.

6. Conclusion

There is a growing consensus that public investment in childcare can advance equity and long-term growth (Devercelli and Beaton-Day, 2020). Multiple studies suggest that early childhood education programs significantly improve children's development outcomes (Britto et al., 2017; Heckman, 2006). Greater stock of human capital due to such programs is likely to contribute to long-term growth (Barro, 2001; Baldacci et al., 2008; Bhattacharyya, 2009). There is also extensive evidence of short-term impacts of childcare through increases in women's labor market engagement. However, women may only be switching from unpaid domestic work to paid work (Halim et al., 2022), which leaves open the question about the productivity gains in the process. Shedding light on the short-term impact is particularly important as governments' policies are often driven by short-term considerations.

Our paper unveils a different – to the best of our knowledge – novel short-term impact of childcare provision, i.e., its potential to trigger short-term economic growth by raising firm productivity. This is conceivable as childcare may enable a better allocation of talent across different occupations, with the home sector as one of the possible occupations (Hsieh et al., 2019).

[8] The number of plants is slightly reduced relative to our usual sample as the skill question constrains the sample to plants already existing in 2006.

Using a combination of a manufacturing survey and data on expansion of preschools in Indonesia, we provide causal evidence of the impact of institutional childcare availability on manufacturing plants' productivity. We do so by employing a triple diff-in-diff estimation comparing the outcomes of plants in sectors that rely on female labor at baseline to different degrees. This allows us to address several key threats to the identification by employing a large array of time varying fixed effects. Our preferred results suggest that an additional preschool per 1,000 children (approximately 1/3 standard deviation increase in preschool density) boosts TFPR by 11 percent for plants with average fraction of women among their employees. The increase is driven by changes in technical efficiency of the firms: we find zero impact on markup and impacts on TFPR and TFPQ are of similar magnitude.

Our results are robust to various changes in specification. In addition to our preferred OLS specification, we measure impacts of preschool expansion on plant productivity in an instrumental variables framework. We also confirm the robustness of the estimates to using different levels of sectoral aggregation, different methods of accounting for missing years in preschool data, and to including a rich set of time-variant control variables. We find suggestive evidence in support of two channels underlying the observed increase in productivity: better allocation of talent leading to improved matching between firms (à la Pissarides, 2000), and increase in job stability, which we measure through a reduction in turnover.

Our study offers the first evidence that provision of childcare may enable governments to reap short-term economic benefits through higher firm productivity. To what extent this effect could hold beyond the specific Indonesian context remains an open question. At the beginning of the study period only 49 percent of Indonesian women worked, and a large fraction of these women were highly educated. Moreover, Indonesian firms, more than their counterparts in the neighboring countries, suffered from low access to qualified personnel. Thus, removing childcare constraints to women's labor force participation resulted in an influx of educated workers in the labor market. In a context where women who are out of the labor force have lower education levels, investment in childcare may generate lower immediate benefits. As such, more research is needed to test for the external validity of the impacts of access to childcare on firm productivity.

It would also be important to continue studying the long-run impacts of ECED programs in contexts like Indonesia. These are complementary to the childcare dimension of such programs, which we have focused on in this paper. Impact evaluations of some Indonesian ECED programs (Chang et al., 2006) point to the potential of preschools to also generate a boost to human capital development, positively impacting long-term growth.

References

Andresen, Martin, and Tarjei Havnes. “Child Care, Parental Labor Supply and Tax Revenue.” SSRN Electronic Journal, 2018. DOI.org (Crossref), https://doi.org/10.2139/ssrn.3193330.

Auger, Anamarie, George Farkas, Margaret Burchinal, Greg Duncan, and Deborah Vandell, “Preschool Center Care Quality Effects on Academic Achievement: An Instrumental Variables Analysis,” Developmental psychology, 12 2014, 50, 2559–71.

Baker, Michael, Jonathan Gruber and Kevin Milligan (2008) Universal child care, maternal labor supply, and family well-being. Journal of Political Economy, 116 (4) (2008), pp. 709-745

Baldacci, Emanuele, Benedict Clements, Sanjeev Gupta, and Qiang Cui. “Social Spending, Human Capital, and Growth in Developing Countries.” World Development 36, no. 8 (August 2008): 1317–41. https://doi.org/10.1016/j.worlddev.2007.08.003.

Barro, Robert J. “Human Capital and Growth.” American Economic Review 91, no. 2 (May 1, 2001): 12–17. https://doi.org/10.1257/aer.91.2.12.

Barros, Ricardo P., Pedro Olinto, Trine Lunde, and Mirela Carvalho. 2011. “The Impact of Access to Free Childcare on Women’s Labor Market Outcomes: Evidence from a Randomized Trial in Low-Income Neighborhoods of Rio de Janeiro.” World Bank Policy Report, Washington, DC.

Beerli, Andreas, Jan Ruffner, Michael Siegenthaler, and Giovanni Peri. “The Abolition of Immigration Restrictions and the Performance of Firms and Workers: Evidence from Switzerland.” American Economic Review 111, no. 3 (March 1, 2021): 976–1012. https://doi.org/10.1257/aer.20181779.

Bhattacharyya, Sambit. “Unbundled Institutions, Human Capital and Growth.” Journal of Comparative Economics 37, no. 1 (March 2009): 106–20. https://doi.org/10.1016/j.jce.2008.08.001.

Bloom, David, David Canning, Günther Fink, and Jocelyn E. Finlay (2009). “Fertility, Female Labor Force Participation, and the Demographic Dividend.” Journal of Economic Growth, 14 (2): 79-101.

Britto, Pia R, Stephen J Lye, Kerrie Proulx, Aisha K Yousafzai, Stephen G Matthews, Tyler Vaivada, Rafael Perez-Escamilla, et al. “Nurturing Care: Promoting Early Childhood Development.” The Lancet 389, no. 10064 (January 2017): 91–102. https://doi.org/10.1016/S0140-6736(16)31390-3.

Brown, James N. 1989."Why do wages increase with tenure? On-the-job training and life-cycle wage growth observed within firms." *The American Economic Review*: 971-991.

Brown, Charles and James Medoff (1978). “Trade Unions in the Production Process”. Journal of Political Economy 86 (3), pp. 355–78.

Calì, Massimiliano, Nicola Cantore, Leonardo Iacovone, Mariana Pereira-López, and Giorgio Presidente. 2022. Too much energy: The perverse effect of low fuel prices on firms." *Journal of Environmental Economics and Management* 111 (2022): 102587.

Cali, Massimiliano; Le Moglie, Marco; Presidente, Giorgio. 2021. "Gain without Pain? Non-Tariff Measures, Plants' Productivity and Markups." Policy Research Working Paper No. 9654. World Bank, Washington, DC.

Cascio, Elizabeth U., "Does Universal Preschool Hit the Target? Program Access and Preschool Impacts," Journal of Human Resources, Forthcoming.

Chang, Mae Chu Chen, Dandan Cibulskis, Richard Dunkelberg, Erika Iskandar, Susiana Josodipoero, Ratna I. Namasivayam, Puspavali Naudeau, Sophie Soekatri, Moesijanti Warouw, Elsa Febianna. 2006. "Early childhood education and development in Indonesia : an investment for a better life" Working paper series ; no. 2006-2 Washington, D.C. : World Bank Group.

Clark, Shelley, Caroline W. Kabiru, Sonia Laszlo, and Stella Muthuri. 2019. "The Impact of Childcare on Poor Urban Women's Economic Empowerment in Africa." Demography 56 (4): 1247–72. https://doi.org/10.1007/s13524-019-00793-3.

Del Boca, Daniela. 2015. "Child Care Arrangements and Labor Supply." Working Paper IDB-WP-569. IDB Working Paper Series. https://www.econstor.eu/handle/10419/115499.

Devercelli, Amanda E., and Frances Beaton-Day. 2020. Better Jobs and Brighter Futures: Investing in Childcare to Build Human Capital. World Bank. https://doi.org/10.1596/35062.

Garces, Eliana, Duncan Thomas, and Janet Currie. "Longer-Term Effects of Head Start." American Economic Review 92, no. 4 (August 1, 2002): 999–1012. https://doi.org/10.1257/00028280260344560.

Garcia, Jorge Luis, et al. "Quantifying the Life-Cycle Benefits of a Prototypical Early Childhood Program." Journal of Political Economy, 2020.

Gertler, Paul, et al. "Labor Market Returns to an Early Childhood Stimulation Intervention in Jamaica." Science, vol. 344, no. 6187, May 2014, pp. 998–1001. DOI.org (Crossref), https://doi.org/10.1126/science.1251178.

Giesing, Yvonne and Laurentsyeva, Nadzeya, Firms Left Behind: Emigration and Firm Productivity (January 15, 2018). CESifo Working Paper Series No. 6815, Available at SSRN: https://ssrn.com/abstract=3126663 or http://dx.doi.org/10.2139/ssrn.3126663

Grinza, Elena, and François Rycx. "The impact of sickness absenteeism on firm productivity: new evidence from Belgian matched employer–employee panel data." *Industrial Relations: A Journal of Economy and Society* 59, no. 1 (2020): 150-194.

Halim, Daniel, Hillary C. Johnson, and Elizaveta Perova. 2022. *Preschool Availability and Women's Employment: Evidence from Indonesia*. Economic Development and Cultural Change, forthcoming. https://doi.org/10.1596/1813-9450-8915.

Halim,Daniel, Elizaveta Perova and Sarah Reynolds. 2021. “Childcare and Mothers’ Labor Market Outcomes in Lower- and Middle-Income Countries”. Policy Research working paper WPS 9828 Washington, D.C. : World Bank Group.

Harper, Sam, Nichole Austin, and Arijit Nandi. 2017. “Daycare and Women’s Health, Social, and Economic Outcomes in Low- and Middle-Income Countries: Systematic Review and Evidence Synthesis.” OSF Preprints. https://doi.org/10.31219/osf.io/xm8g6.

Heckman, James J. “Skill Formation and the Economics of Investing in Disadvantaged Children.” Science 312, no. 5782 (June 30, 2006): 1900–1902. https://doi.org/10.1126/science.1128898.

Hegewisch, Ariane, and Janet C. Gornick. 2011. “The Impact of Work-Family Policies on Women’s Employment: A Review of Research from OECD Countries.” *Community, Work & Family* 14 (2): 119–38. https://doi.org/10.1080/13668803.2011.571395.

Herbst, Chris M. 2017. “Universal Child Care, Maternal Employment, and Children’s Long-Run Outcomes: Evidence from the US Lanham Act of 1940,” *Journal of Labor Economics*, 35 (2), 519–564.

Hsieh, Chang-Tai, Erik Hurst, Chad Jones and Peter Klenow. 2019. "The Allocation of Talent and U.S. Economic Growth, *Econometrica*.

Michele Kacmar, K., Martha C. Andrews, David L. Van Rooy, R. Chris Steilberg, and Stephen Cerrone. "Sure everyone can be replaced… but at what cost? Turnover as a predictor of unit-level performance." *Academy of Management journal* 49, no. 1 (2006): 133-144.

Mertens, M. (2019). Micro-mechanisms behind declining labour shares: Market power, production processes, and global competition. Technical report, IWH-CompNet Discussion Papers.

Morrissey, Taryn W. 2017. “Child Care and Parent Labor Force Participation: A Review of the Research Literature.” *Review of Economics of the Household* 15 (1): 1–24. https://doi.org/10.1007/s11150-016-9331-3.

Olivetti, Claudia, and Barbara Petrongolo. 2017. “The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries.” *Journal of Economic Perspectives* 31 (1): 205–30. https://doi.org/10.1257/jep.31.1.205.

Padilla-Romo, María, and Francisco Cabrera-Hernández. 2018. “The Effect of Children’s Time in School on Mothers’ Labor Supply : Evidence from Mexico’s Full-Time Schools Program.” Working Paper Series of the Department of Economics. Haslam College of Business. https://repositorio.minedu.gob.pe/handle/20.500.12799/5962.

Pissarides, Christopher. 2000. Equilibrium Unemployment Theory (Second ed.). Cambridge, MA: MIT Press. 2000. ISBN 978-0-262-16187-9.

Ranganathan, Aruna, and David S. Pedulla. “Work-Family Programs and Nonwork Networks: Within-Group Inequality, Network Activation, and Labor Market Attachment.” Organization Science 32, no. 2 (March 2021): 315–33. https://doi.org/10.1287/orsc.2020.1391.

Roy, Andrew Donald. "Some thoughts on the distribution of earnings." *Oxford economic papers* 3, no. 2 (1951): 135-146.

Schaner, Simone and Das, Smita, Female Labor Force Participation in Asia: Indonesia Country Study (February 10, 2016). Asian Development Bank Economics Working Paper Series No. 474, Available at SSRN: https://ssrn.com/abstract=2737842 or http://dx.doi.org/10.2139/ssrn.2737842

Shafiq, M. Najeeb, Amanda Devercelli, and Alexandria Valerio. "Are There Long-Term Benefits from Early Childhood Education in Low- and Middle-Income Countries?" SSRN Electronic Journal, 2018. https://doi.org/10.2139/ssrn.3270603.

Shaw, Jason D. (2011). "Turnover rates and organizational performance: Review, critique, and research agenda". Organizational Psychology Review 1(3), pp. 187–213.

Strober, Myra H. (1990). "Challenges to Human Capital Theory: Implications for HR Managers". Industrial Relations Spring.

Stock, James H., Jonathan H. Wright, and Motohiro Yogo. "A survey of weak instruments and weak identification in generalized method of moments." Journal of Business & Economic Statistics 20, no. 4 (2002): 518-529.

Yanadori, Yoshio and Takao Kato (2007). "Average employee tenure, voluntary turnover ratio, and labour productivity: evidence from Japanese firms". The International Journal of Human Resource Management 18, (10).

Preschools and firm productivity – Tables

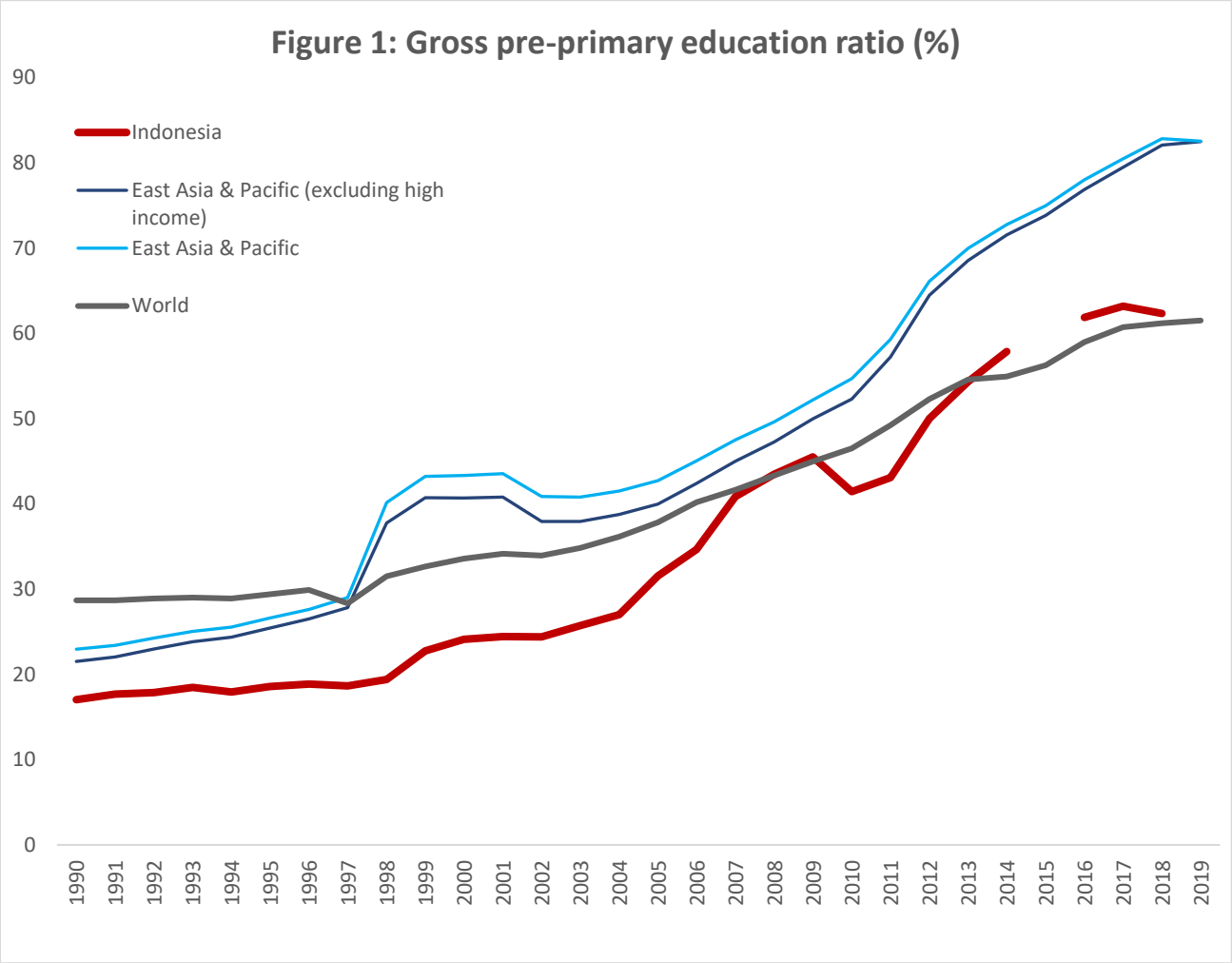
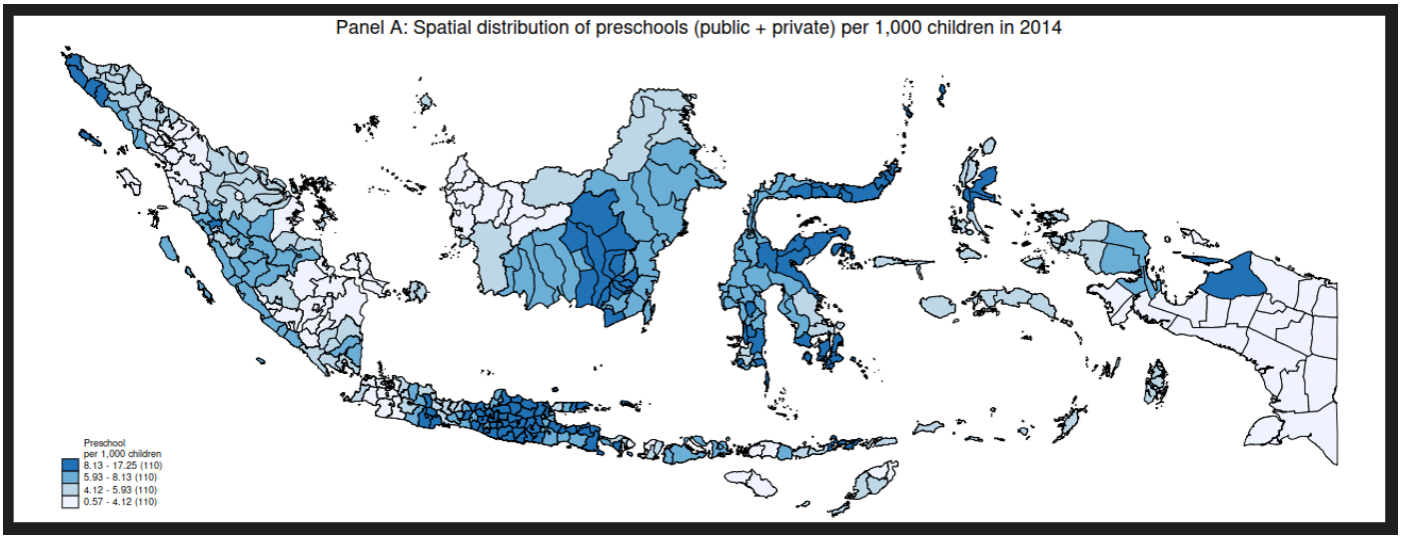


Figure 2


Panel A: Spatial distribution of preschools (public + private) per 1,000 children in 2014
Preschool
per 1,000 children
8.13 - 17.25 (110)
5.93 - 8.13 (110)
4.12 - 5.93 (110)
0.57 - 4.12 (110)

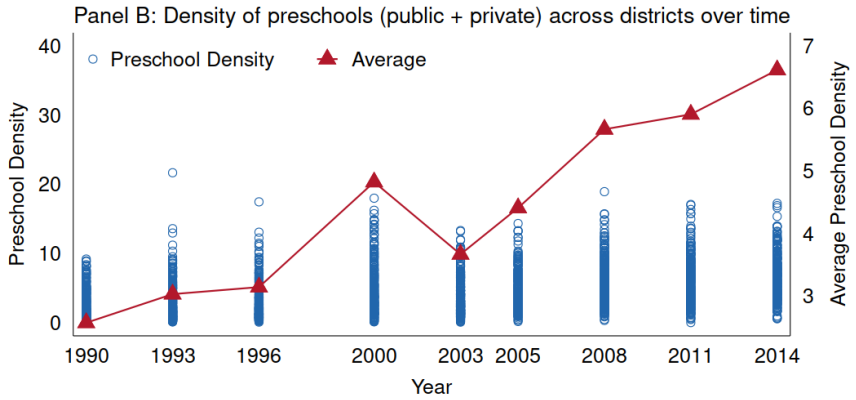

Panel B: Density of preschools (public + private) across districts over time
Preschool Density
Average
Preschool Density
Average Preschool Density
Year
0
10
20
30
40
3
4
5
6
7
1990
1993
1996
2000
2003
2005
2008
2011
2014

Table 1: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Manufacturing survey					
Log TFPQ	167,175	-0.185	11.387	-53.659	32.061
Log TFPR	153,361	8.148	11.583	-44.460	49.420
Log Markup	156,589	8.680	1.503	-3.038	26.457
Share of female workers	31,922	0.410	0.296	0	1
Log Average Income	264,108	9.130	0.995	4.605	14.686
Log Female Worker	253,836	2.99	1.54	0.00	10.87
Total workers	265,753	199.509	762.334	20.000	57383.980
Skill gap	228,697	.0127	0.112	0	1
Log Raw Materials	265,950	14.180	3.7212	-0.157	25.096
Share of Foreign Ownership	265,972	7.7646	25.430	0	100
PODES					
Preschool Density (number of preschools per 1,000 children of eligible age, by district)	247,137	5.626	3.015	0.029	18.934
SAKERNAS					
Employment in manufacturing	5,257,459	0.178	0.383	0	1
Employment in services	5,257,459	0.057	0.231	0	1
Wage employment	5,257,459	0.246	0.431	0	1
Female	5,257,459	0.503	0.500	0	1
Education	5,257,459	7.248	4.677	0	18
Experience	4,970,262	28.659	15.75	3	67
INDO-DAPOER					
District Real GDP	3,757	2.12E+07	3.65E+07	350567.9	3.33E+08
Total Length of Road, km	3,692	146.457	163.340	0.178	1555.683
Average national exam score: Lower secondary level (out of 100)	3,757	60.183	7.836	42.671	84.913
Total government expenditure (realization), in IDR	3,691	9.53E+11	1.05E+12	4.78E+09	2.54E+13

Table 2: Testing for parallel trends

	(1) TFPR	(2) TFPQ	(3) Markup	(4) TFPR	(5) TFPQ	(6) Markup
Preschool density change x Sector exposure x Post Dummy	-0.002	0.011	-0.013	-0.155*	-0.223**	0.068*
	(0.057)	(0.054)	(0.011)	(0.089)	(0.092)	(0.039)
Observations	31,242	31,242	31,242	31,377	31,377	31,377
R-squared	0.962	0.965	0.846	0.962	0.965	0.846
Post dummy cutoff	2001	2001	2001	2002	2002	2002

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table shows firm-level regression between dependent variables (TFPQ, TFPR, and Markup) with triple interaction of change in preschool density between 2002-2014, sector-level average of share of female worker in 2002, and a dummy variable with value 1 if year of observation is equal to or greater than the cutoff point. Controls include district-year fixed-effects, sector-year fixed-effects, plant fixed-effect, raw materials and share of foreign ownership. The regressions are done with manufacturing survey data between 2000 and 2002. Standard errors are clustered at district level.

Table 3: Testing for parallel trends

	(1) TFPR	(2) TFPQ	(3) Markup
Preschool density change x Sect exp	-0.155	-0.222*	0.067
	(0.125)	(0.129)	(0.057)
Observations	9,168	9,168	9,168
R-squared	0.109	0.108	0.124
District FE	Yes	Yes	Yes
Sector FE	5-digit	5-digit	5-digit

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, *
p<0.1

Table shows firm-level regression between dependent variables (TFPR, TFPQ, and Markup) and an interaction variable consisting change in preschool density between 2002-2014 interacted with sector-level average share of female workers in 2002. Controls include district fixed-effects and sector fixed-effects. We use manufacturing survey data in 2000 and 2002 for the regression. Standard errors are clustered at district level.

Table 4: Productivity impact of preschools - Baseline estimates

	(1) TFPR	(2) TFPQ	(3) Markup
Preschool Density x Share of Female Worker	0.259***	0.244***	0.0163
	(0.0854)	(0.0810)	(0.0234)
Observations	140,855	153,808	143,911
R-squared	0.945	0.948	0.811
Control	Yes	Yes	Yes
District-Time FE	Yes	Yes	Yes
Sector-Time FE	5 digit	5 digit	5 digit
Plant FE	Yes	Yes	Yes
Estimation Method	OLS Reduced Form	OLS Reduced Form	OLS Reduced Form

*Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1*

Table shows firm-level regression between dependent variables (TFPR, TFPQ, Markup) and an interaction variable consisting preschool density and 5-digit sector-level average of share of female worker in 2002. Controls include plant fixed-effects, district-time fixed-effects, sector-time fixed-effects, raw materials, and share of foreign ownership. Regressions done using manufacturing survey data spanning 2002-2014. Standard errors are clustered at district level.

Table 5: Productivity impact of preschools - Instrumental variables estimates

	(1) TFPR	(2) TFPQ	(3) Markup
Preschool Density x Share of Female Worker	0.261***	0.243***	0.0164
	(0.0867)	(0.0815)	(0.0236)
Observations	140,855	153,808	143,911
R-squared	0.010	0.000	0.123
Control	Yes	Yes	Yes
District-Time FE	Yes	Yes	Yes
Sector-Time FE	5 digit	5 digit	5 digit
Plant FE	Yes	Yes	Yes
First stage F-stat	1292	1315	1282
FemShr Instr.	5-digit	5-digit	5-digit

*Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1*

Table shows firm-level regression between dependent variables (TFPR, TFPQ, Markup) and an interaction variable consisting preschool density and plant-level share of female worker in 2002, instrumented with an interaction between preschool density and 5-digit sector-level average of share of female worker in 2002. Controls include plant fixed-effects, district-time fixed-effects, sector-time fixed-effects, raw materials, and share of foreign ownership. We use manufacturing survey data spanning 2002-2014. Standard errors are clustered at district level.

Table 6: Baseline results are robust to using different level of aggregation

	TFPR	TFPQ	Markup	TFPR	TFPQ	Markup
	(I)	(II)	(III)	(IV)	(V)	(VI)
Preschool Density x Share of Female Worker	0.280***	0.228**	0.0158	0.302***	0.236**	0.0170
	(0.0926)	(0.0908)	(0.0267)	(0.102)	(0.0948)	(0.0289)
Observations	141,250	154,207	144,302	141,250	154,207	144,302
R-squared	0.942	0.945	0.803	0.011	0.000	0.123
Control	Yes	Yes	Yes	Yes	Yes	Yes
District-Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Sector-Time FE	3 digit	3 digit	3 digit	3 digit	3 digit	3 digit
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-stat	-	-	-	245	314	232
Method	OLS Reduced Form	OLS Reduced Form	OLS Reduced Form	IV	IV	IV

Table shows firm-level regression between dependent variables (TFPR, TFPQ, Markup) and an interaction variable consisting preschool density and plant-level share of female worker in 2002. Regressions are done either using ordinary least squares or instrumental variable approach, where we use interaction between preschool density and 3-digit sector-level average of share of female worker in 2002 as instruments. Controls include plant fixed-effects, district-time fixed-effects, 3-digit sector-time fixed-effects, raw materials, and share of foreign ownership. We use manufacturing survey data spanning 2002-2014. Standard errors are clustered at district level.

Table 7: Baseline results are robust to using different strategies to deal with missing preschool years

	(1) TFPR	(2) TFPQ	(3) Markup	(4) TFPR	(5) TFPQ	(6) Markup
Preschool Density x Share of Female Worker	0.203*	0.165*	0.0360	0.321***	0.264***	0.0515*
	(0.108)	(0.0962)	(0.0359)	(0.0982)	(0.0913)	(0.0288)
Observations	46,598	51,203	47,697	140,460	153,377	143,504
R-squared	0.954	0.955	0.838	0.945	0.948	0.811
Control	Yes	Yes	Yes	Yes	Yes	Yes
District-Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Sector-Time FE	5 digit	5 digit	5 digit	5 digit	5 digit	5 digit
Plant FE	Yes	Yes	Yes	Yes	Yes	Yes
Missing method	PODES years only			Linear projection		

*Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1*

Table shows firm-level regression between dependent variables (TFPR, TFPQ, Markup) and an interaction variable consisting preschool density and 5-digit sector-level average of share of female worker in 2002. Preschool density is calculated either directly from PODES data that is only available every three years, or we take linear projections to impute the values for the years when the data is not directly observed. Controls include plant fixed-effects, district-time fixed-effects, sector-time fixed-effects, raw materials, and share of foreign ownership. Regressions done using manufacturing survey data spanning 2002-2014. Standard errors are clustered at district level.

Table 8: Checking that the results are not driven by the correlation between preschool expansion and other changes that may affect female dominated plants differentially

	(1)	(2)	(3)	(4)	(5)	(6)
	TFPR	TFPQ	TFPR	TFPQ	TFPR	TFPQ
Preschool Density x Share of Female Worker	0.226***	0.216***	0.259***	0.243***	0.267***	0.250***
	(0.0863)	(0.0817)	(0.0854)	(0.0810)	(0.0856)	(0.0812)
Control x Share of Female Worker	-0.533	-0.485	0.239***	0.216***	-0.0043	-0.0045
	(0.341)	(0.339)	(0.0820)	(0.0722)	(0.0034)	(0.0029)
Control in interaction	gov. expenditure		exam scores		infrastructure	
Observations	131,561	144,314	140,043	152,940	140,006	152,903
R-squared	0.947	0.950	0.945	0.948	0.945	0.948
Control	Yes	Yes	Yes	Yes	Yes	Yes
District-Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Sector-Time FE	5 digit	5 digit	5 digit	5 digit	5 digit	5 digit
Plant FE	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table shows regression between dependent variables (TFPR and TFPQ) and two interaction variables consisting preschool density interacted with 5-digit sector-level average of share of female worker in 2002, and district-level covariates (government expenditure, exam scores, and infrastructure) interacted with 5-digit sector-level average of share of female worker in 2002. Controls include plant fixed-effects, district-time fixed-effects, sector-time fixed-effects, raw materials, and share of foreign ownership. Regressions done using manufacturing survey data spanning 2002-2014. Standard errors are clustered at district level.

Table 9: Preschools increase the share of female workers

	(1)	(2)	(3)	(4)
	Shr Fem Workers	Shr Fem Workers	Fem Workers	Fem Workers
Preschool Density x Share of Female Worker	0.0097***	0.0094***	0.0239	0.0089
	(0.0025)	(0.0029)	(0.0148)	(0.0171)
Observations	244,158	140,855	232,658	134,586
R-squared	0.896	0.907	0.890	0.916
Control	Yes	Yes	Yes	Yes
District-Time FE	Yes	Yes	Yes	Yes
Sector-Time FE	5 digit	5 digit	5 digit	5 digit
Plant FE	Yes	Yes	Yes	Yes
Sample	Full	TFP	Full	TFP

*Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1*

Table shows firm-level regression between dependent variables (plant-level share of female workers, plant-level number of female workers) and an interaction variable consisting change in preschool density between 2002-2014 interacted with sector-level average share of female workers in 2002. Controls include district fixed-effects and sector fixed-effects. We use manufacturing survey data in 2000 and 2002 for the regression. We either include full sample or use the subset of the data where we have managed to construct the TFP variables. Standard errors are clustered at district level.

Table 10: Preschools increase employment participation of female across sectors, particularly for women with young kids

	(1)	(2)	(3)	(4)	(5)	(6)
	Empl Mauf	Empl Serv	Wage Empl	Empl Mauf	Empl Serv	Wage Empl
Female	-0.0199***	-0.0767***	-0.121***			
	(0.00153)	(0.00300)	(0.00321)			
Preschool x Female	0.907***	5.634***	3.992***			
	(0.208)	(0.371)	(0.403)			
Kids Below 10				-0.0108***	-0.0436***	-0.0463***
				(0.00171)	(0.00250)	(0.00291)
Preschool x Kids Below 10				1.174***	2.438***	1.617***
				(0.232)	(0.336)	(0.361)
Observations	5,208,237	5,208,237	5,208,237	1,214,623	1,214,623	1,214,623
R-squared	0.050	0.076	0.094	0.059	0.057	0.066
District-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	Female	Female	Female

*Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1*

Table shows individual-level regressions using labor force survey data spanning 2007-2014. Dependent variables are dummy variables indicating if the respondent worked in manufacturing sector, services sector, or is employed as wage employee in any sector. In column 1-3, regressions are done with the universe of working age population in the sample period. In column 4-6, regressions are done for only female working age population that answers to the question on whether the female has any kids at the time of the data collection. Standard errors are clustered at district-year level.

Table 11: Preschools reduce the women's turnover rate in manufacturing, particularly for women with young kids

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent			*Change job in past 12 months*			
Sample	Employed in Mfg	Employed in Services	Employed overall	Fem emp Mfg	Fem emp serv	Fem. Emp. overall
Female	0.0159***	-0.00514***	-0.00387***			
	(0.00350)	(0.00101)	(0.000771)			
Preschool x Female	-1.380***	-0.0361	-0.0125			
	(0.449)	(0.155)	(0.128)			
Kids Below 10				0.0127**	-0.0123***	-0.00482***
				(0.00571)	(0.00223)	(0.00135)
Preschool x Kids Below 10				-3.036***	-0.554	-0.537**
				(0.788)	(0.346)	(0.212)
Observations	310,254	1,312,494	3,390,305	63,000	255,929	605,285
R-squared	0.044	0.020	0.029	0.071	0.031	0.044
District-Year FE	Yes	Yes	Yes	Yes	Yes	Yes

*Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1*

Table shows individual-level regressions using labor force survey data spanning 2007-2014. Dependent variables are dummy variables indicating if the respondent reports on quitting job within 12 months prior to the data collection period, either from manufacturing sector, services sector, or any sector. In column 1-3, regressions were done with the universe of female working age population that reported on quitting jobs, or female respondents that were employed at the time of the data collection. In column 4-6, regressions were done for only employed female working respondents that answer to the question on whether the female has any kids at the time of the data collection.

Table 12: Preschools increase workers' incomes particularly in plants with initial more stringent skills' gap

	(1)	(2)	(3)	(4)
Dependent variable		Workers' average income		
FemShr x Pre-Sch.	0.0359**	0.0348*	0.0411**	0.0470**
	(0.0160)	(0.0208)	(0.0163)	(0.0213)
FemShr x Pre-Sch. x Skill Gap			0.0622**	0.0672
			(0.0276)	(0.0429)
Observations	242,350	139,791	211,775	121,258
R-squared	0.689	0.716	0.685	0.715
Control	Yes	Yes	Yes	Yes
District-Time FE	Yes	Yes	Yes	Yes
Sector-Time FE	5 digit	5 digit	5 digit	5 digit
Plant FE	Yes	Yes	Yes	Yes
Sample	Full	TFP	Full	TFP

*Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1*

Table shows firm-level regression between log of workers average income and an interaction variable consisting preschool density and 5-digit sector-level average of share of female worker in 2002. Other covariate included is the triple interaction between preschool density, 5-digit sector-level average of share of female worker in 2002, and a dummy variable equals to 1 if the firm reported lack of skill as one the firm's primary constraints in 2006 survey. Controls include plant fixed-effects, district-time fixed-effects, sector-time fixed-effects, raw materials, and share of foreign ownership. Regressions done using manufacturing survey data spanning 2002-2014. Standard errors are clustered at district level.

Appendix I: Estimation of performance measures

We assume that in each period t , plant j produces output Q_{jt} with the following production function:

$$Q_{jt} = \min\{\gamma_e E_{jt}, F(K_{jt}, L_{jt}, M_{jt}) \cdot \Omega_{jt}\} \qquad \text{(A1)}$$

where E_{jt} is energy consumption, K_{jt} the capital stock, L_{jt} labor and $M_{jt} = \{M_{jt}^d, M_{jt}^i\}$ are domestic and imported raw materials, respectively. The term Ω_{jt} represents Hicks-neutral productivity. The production function (A1) is a structural value added specification (De Loecker & Scott, 2016) in which capital, labor and materials are allowed to be characterized by some degree of substitution and energy is a perfect complement to the combination of the other inputs. Most existing structural value added specifications include raw materials (possibly including energy) in the first argument of the $\min$ operator, therefore considering them proportional to output. Proportionality is equivalent to assuming that plants never change the materials' intensity of production. However, that would not be appropriate in our framework, for two reasons. First, as emphasized by the literature on per-shipment costs, the latter might induce shipment delays or trigger substitution between imported and domestic materials. In turn, changes in the mix of materials are likely to affect the whole production process, which involves adjustment in capital and labor. The second reason is that Indonesia is often characterized by poor transportation infrastructure, which, combined with a high geographical heterogeneity, can result in road closures, disruption in naval shipments or technical failure of the material supplier's transportation equipment (Cali, 2022). Therefore, it seems appropriate to allow for some degree of adjustment costs for raw materials and consider energy a more flexible input.[9]

Given (A1), a profit maximizing plant would set:

$$Q_{jt} = \gamma_e E_{jt} = F(K_{jt}, L_{jt}, M_{jt}) \cdot \Omega_{jt} \qquad \text{(A2)}$$

Recovering Production Function Parameters

This section describes the estimator used to recover the production function parameters. As in Mertens (2019), we exploit product-level information on quantities and prices to construct plant-level deflators which we use to deflate revenues. Unlike most of the existing literature, however, we are also able to construct plant-level inputs' deflators by exploiting information on price and quantities of each raw material used in production by the plant. This is a strong advantage over previous studies such as Mertens (2019) and De Loecker, et al. (2016) that do not have plant-level materials' deflators. Using plant-specific price indices allows us to overcome the bias that would otherwise arise using revenues and materials' expenditure deflated by industry price indices.[10]

Given that we observe quantities for products and materials, an alternative approach could have been using output and input quantity directly in our estimation routine. However, following such a strategy would pose an aggregation issue due to plants often reporting the same products and inputs in different units, which would introduce considerable noise. Moreover, since quantity output and input data are

[9] Notice that in absence of such adjustment costs, our specification would still be valid.

[10] This would be equivalent to assuming that all plants have the same product within an industry, or that they use materials of identical quality.

available for only 63% of the plants, using plant-specific deflators to derive quasi-quantities allows us to maximize the size of the sample.

One problem we still face, common to all existing studies estimating production functions, is that we do not observe physical capital. Since different plants might use capital assets of different qualities, using expenditure to estimate productivity would result in biased estimates. We address the issue by using asset type-specific price indices to deflate the value of the capital stock[11] and estimating the production function parameters by 2-digit industries separately, as the homogeneity of capital equipment within such industries gives less scope for capital price variation.[12]

In our empirical application, we use a flexible trans-log specification $f(\cdot)$. We estimate the logged version of production function (A2):

$$q_{jt} = f\left(k_{jt}, l_{jt}, m_{jt}^{d}, m_{jt}^{i}; \boldsymbol{\beta}\right) + \omega_{jt} + \varepsilon_{jt} \qquad \text{(A3)}$$

Recall that ω_{jt} represents the log of Hicks-neutral productivity, which is known by plants' managers but not by us. The variable ε_{jt} is an i.i.d. error term that captures factors such as measurement errors.

We are interested in estimating the vector of the production function parameters $\boldsymbol{\beta}$. To recover unbiased and consistent estimates of firms' production function (A3), we need to address the well-known simultaneity problem deriving from the fact that ω_{jt} is correlated to labor and materials' input (but not to capital, which is chosen one period ahead). We build on the methodology of Ackerberg et al. (2015). In particular, we make the following timing assumptions concerning inputs' decisions: i) capital k_{jt} is chosen at t -1; ii) I_{jt}, m_{jt}^{d}, and m_{jt}^{i} are chosen in $t-b$ after observing ω_{jt}, and iii) energy e_{jt} is chosen at $t-a$, with $1 < b < a$.

Our timing assumptions differ from most of the existing literature, which typically assumes materials to be completely flexible inputs. We depart from that assumption in order to allow for materials' adjustment costs. We use our timing assumption to specify an energy demand function,

$$e_{jt} = \tilde{h}\left(\omega_{jt}, k_{jt}, l_{jt}, m_{jt}^{d}, m_{jt}^{i}, \boldsymbol{\theta_{jt}}\right) \qquad \text{(A4)}$$

The vector $\boldsymbol{\theta_{jt}}$ includes variables that affect plant-level demand for energy. Cali et al. (2022) show that due a variegated landscape differences in the quality of infrastructures, the cost of energy distribution is heterogeneous across Indonesian regions. Therefore, we include a full set of location dummies in $\boldsymbol{\theta_{jt}}$. We also include year dummies which capture time varying factors common to all plants in a given industry.

[11] We obtain wholesale price indices for different asset-types from the Indonesian Statistical Office (BPS).
[12] We do not have a sufficient number of observations to robustly estimate the parameters by three or more digits industries.

Assuming that the energy consumption function of the plant, $\tilde{h}$ is monotonically increasing and invertible in ω, we obtain a control function that proxies for unobserved productivity,

$$\omega_{jt} = h\left(e_{jt}, k_{jt}, l_{jt}, m_{jt}^{d}, m_{jt}^{i}, \boldsymbol{\theta_{jt}}\right) \quad \text{(A5)}$$

where $h \equiv \widetilde{h^{-1}}$. One additional advantage of our approach is that in our data, we directly observe the quantity of electricity consumed.[13] That makes our estimator less subject to bias. To see this, consider an exogenous shock to energy prices. If we were to use energy expenditures rather than actual quantities, depending on the elasticity of electricity consumption to energy prices, the increase in prices might result in an increase of energy expenditures and lead us to erroneously conclude that unobserved productivity increased based on (A5).

Adding $h(\cdot)$ to (A5), we get

$$q_{jt} = f\left(k_{jt}, l_{jt}, m_{jt}^{d}, m_{jt}^{i}; \boldsymbol{\beta}\right) + h\left(e_{jt}, k_{jt}, l_{jt} m_{jt}^{d} m_{jt}^{i}, \boldsymbol{\theta_{jt}}\right) + \varepsilon_{jt} \quad \text{(A6)}$$

We follow Ackerberg et al. (2015) by approximating the right-hand-side of (A5) with a third-order polynomial in all its elements, except for the elements of $\boldsymbol{\theta_{jt}}$, which we enter linearly.[14] From the first stage, we obtain expected output $\widehat{q_{jt}}$ and the residuals $\widehat{\varepsilon_{jt}}$.[15]The next step is specifying a law of motion for productivity ω_{jt}. We assume that ω_{jt} follows a Markov process that can be shifted by the plant managers' action:

$$\omega_{jt} = g\left(\omega_{j,t-1}, \boldsymbol{\Gamma_{j,t-1}}\right) + \xi_{jt} \quad \text{(A7)}$$

In (A7), ξ_{jt} denotes the innovation to productivity and the vector $\boldsymbol{\Gamma}$ includes variables controlled by plants' managers that influence the expected future value of productivity and state variables which determine differences in productivity dynamics across plants.[16] In our framework, these are dummy variables for sector fixed effects.

Current expected productivity is then expressed as a function of the data and parameters,

$$\omega(\boldsymbol{\beta})_{jt} = \widehat{q_{jt}} - f\left(k_{jt}, l_{jt}, m_{jt}^{d}, m_{jt}^{i}; \boldsymbol{\beta}\right)$$

[13] For energy types different from electricity we convert in KhW equivalents using standard conversion factors.
[14] This approach is similar to Mertens (2019).
[15] It should be noticed that in the first stage, none of the production function parameters is identified, because they enter both $f(\cdot)$ and $h(\cdot)$.
[16] For instance, in De Loecker et al. (2016), who study the impact of trade reforms, $\boldsymbol{\Gamma}$ included export dummies and import tariffs; De Loecker (2007) includes export quotas; Doraszelski (2013) includes R&D expenditure, and Konings (2015) includes measures of workforce training.

To estimate $\boldsymbol{\beta}$, we form moments based on the innovation ξ_{jt} in the law of motion (A7),

$$\xi(\boldsymbol{\beta_{jt}}) = \omega(\boldsymbol{\beta_{jt}}) - E\left[\omega(\boldsymbol{\beta_{jt}})\middle|\omega(\boldsymbol{\beta_{j,t-1}}), \boldsymbol{\Gamma_{j,t-1}}\middle|\right] \quad \text{(A8)}$$

The moments that identify the parameters are:

$$E\left[\xi(\boldsymbol{\beta})_{jt}\, \boldsymbol{M_{jt}}\right] = 0 \qquad \text{(A9)}$$

where the vector $\boldsymbol{M_{jt}}$ includes current capital, lagged domestic and imported materials, lagged labor, and lagged electricity consumption. It should be noticed that materials, labor and electricity consumption are all significantly correlated within plants over time, which justifies their inclusion in (A9) as instruments.[17]

Equipped with the estimated production function parameters, β, we recover TFPQ by taking the residual:

$$TFPQ_{it} \equiv \widehat{\omega_{jt}} = \widehat{q_{jt}} - f\left(k_{jt}, l_{jt}, m^d_{jt}, m^i_{jt}; \boldsymbol{\beta}\right) \quad \text{(A10)}$$

Deriving Markups from Plants' Cost Minimization

Following De Loecker (2012), we derive an expression for plants' markup as a function of the elasticity of $F(\cdot)$ with respect to labor (a static production input) and the plant's revenue cost shares for labor and energy consumption.

Cost minimization with respect to labor, which we consider a static input, implies the following first order condition:[18]

$$\frac{\partial \mathcal{L}_{jt}}{\partial L_{jt}} = W_{jt} - \widetilde{\lambda_{jt}} - \frac{\partial F\left(K_{jt}, L_{jt}, M_{jt}\right) \cdot \omega_{jt}}{\partial L_{jt}} = 0$$

where $\mathcal{L}$ is plant j's Lagrangian, W_{jt} wages and $\widetilde{\lambda_{jt}}$ the Lagrangian multiplier. Rearranging terms and multiplying both sides of the previous equation by $\frac{L_{jt}}{Q_{jt}}$, we obtain:

[17] In particular, in a regression of current on past electricity consumption - our proxy variable for productivity - including plant, 2-digit industry-year and region-year fixed effects, the autoregressive coefficient is 0.17 and significant at 99 percent.

[18] What matters for the consistency of the model is that labor is chosen in each period t, as opposed to capital, which is chosen in $t-1$.

$\frac{\partial F(K_{jt},L_{jt},M_{jt})\cdot\Omega_{jt}}{\partial L_{jt}}\frac{L_{jt}}{Q_{jt}} = \frac{1}{\widetilde{\lambda_{jt}}}\frac{W_{jt}}{Q_{jt}}$

Following De Loecker (2012), we define the plant's markup over the marginal cost of output λ_{jt} as

$$\widetilde{\mu_{jt}} \equiv \frac{P_{jt}}{\widetilde{\lambda_{jt}}}$$

where P_{jt} is the price of output produced by the firm. The previous equation yields an expression of plants' markup depending on the elasticity of output with respect to the variable input, β_l, and the inverse of the revenue share of expenditure on L_{jt}:

$$\widetilde{\mu_{jt}} = \frac{\partial F(K_{jt}, L_{jt}, M_{jt}) \cdot \Omega_{jt}}{\partial L_{jt}} \frac{L_{jt}}{Q_{jt}} \frac{P_{jt}Q_{jt}}{W_{jt}L_{jt}}$$

As discussed in De Loecker (2016), $\widetilde{\lambda_{jt}}$ expresses the marginal cost of an additional unit of $F(\cdot)$, but the total marginal cost of producing output, λ_{jt}, is given by:

$$\lambda_{jt} = \widetilde{\lambda_{jt}} + \frac{P_{Et}}{\gamma_e}$$

where P_{Et} denotes the price of energy. Therefore, the correct expression of plant $j's$ markup is given by:

$$\mu_{jt} = \frac{\partial F(K_{jt},L_{jt},M_{jt})\cdot\Omega_{jt}}{\partial L_{jt}}\frac{L_{jt}}{Q_{jt}}\frac{P_{jt}Q_{jt}}{W_{jt}L_{jt}+P_{Et}E_{jt}} \qquad \text{(A11)}$$

We can use μ_{jt} to decompose revenue-total factor productivity (TFPR):

$$TFPR_{it} \equiv \frac{P_{it}Q_{it}}{\lambda_{it}F(K_{it},L_{it})} = \mu_{it} \times TFPR_{it} \qquad \text{(A12)}$$

Finally, we use the estimated production function parameters to compute (A11) and (A12). In our application, we drop the top and bottom 1 percent of the TFPQ and markups' estimates in order to avoid outliers driving our findings, although results are consistent if we do not trim the distributions.