THE 1st INTERNATIONAL CONFERENCE ON DATA SCIENCE AND OFFICIAL STATISTICS
ICDSOS
"Strengthening the Contribution of Data Science and Official Statistics to the Society in the Distruption Era"
2021

Does Palapa Ring Project Infrastructure Bridging Connectivity and Economic Activity?

R Eschachasthi[1], T Purwa[2] and D G Cendekia[3]

[1] Badan Pusat Statistik (BPS)-Statistics Indonesia, Jakarta, Indonesia
[2] Badan Pusat Statistik (BPS)-Statistics Indonesia of Bali Province, Indonesia
[3] Badan Pusat Statistik (BPS)-Statistics Indonesia of Kalimantan Selatan Province, Indonesia

*Corresponding author's e-mail: diyang.dgc@bps.go.id

Abstract. This study examines the impact of existence of the Indonesian Palapa Ring Project (PRP) infrastructure on connectivity and economic activities in 46 districts in the West, Central, and East package of PRP in 2015-2020. Connectivity is an internet activity that measured by using percentage of internet use and economic activity is measured by using Gross Regional Domestic Product (GRDP). The fixed effect staggered difference-in-difference is utilized to analyze the panel data obtained from Badan Pusat Statistik (BPS)-Statistics Indonesia. An examination of parallel trend assumptions, robustness check, and heterogeneity analysis are also presented. The results show that PRP infrastructure has a positive and significant impact on connectivity; yet has no significant effect on economic activity. In response to the findings, the policy should be designed by intensifying coverage and quality of the internet; proliferating Information Communication Technology (ICT) facilities in rural areas; and expanding education and digital literacy programs.

1. Introduction

Since digital technology has thrived in the last decade, there are just a few Indonesians that are connected to the internet [1]. In the same period, the number of internet users in neighboring countries, i.e., Thailand, Philippines, and Vietnam increased twice or more than twice of Indonesian internet users percentage.

Nowadays, the internet has grown rapidly and created new opportunities [2]. It plays a prominent role in the Covid-19 pandemic era where people were forced to be more engaged digitally either in advanced or less developed regions. [3] declared that Indonesia's digital economy increased as the second fastest-growing among Southeast Asian countries. Furthermore, the digital economy in 2025 is predicted to bring nearly more than four times the value-added [4]. However, the digital benefit on economic growth will occur for those who can overcome the unbalance of internet users because of geographic characteristics [5]. To meet the challenges, providing better internet access through Palapa Ring Project (PRP) infrastructure is one of Indonesia's strategic plans to narrow the gaps of geographic digital.

The PRP infrastructure constructs submarine and inland optical fiber internet cables in the underdeveloped region or District/Municipality to link the archipelagos to digital connectivity using technological complexity [6]. Famous with the name of "Sky Highway", PRP was categorized into three packages, i.e, West, Central, and East that include 57 non-commercial districts.

On the impact of various Information Communication Technology (ICT) infrastructures, a large body of literature appears various results. [7] showed that broadband infrastructure establishment and penetration tend to increase economic growth in OECD countries. Even though broadband expansion throughout remote regions in the United States (US) caused increasing in local employment rate and average salary, these impacts do not represent that infrastructure expansion improves economic activities [8]. [9] found that economic growth will accelerate the number of social products, which will lead to an increase in living standards and the wellbeing of the population. Meanwhile, by using fuzzy regression discontinuity analysis [10] found no evidence of increasing the United Kingdom (UK) business performance that resulted from the existence of high-speed transfer broadband. Even so, this study showed that improving internet adoption depend on better access to infrastructure.

Various studies have also been conducted to observe the relationship between submarine internet cable on employment and economic growth, [11]–[13]. Meanwhile, studies related to the effect of PRP on human development through a qualitative approach was conducted by [14], while the impact of PRP on the social and economic competitiveness index was done by [15].

Considering the PRP infrastructure developed by using high investment with laying fiber-optic network in the non-commercial region, there is limited policy evaluation about PRP could deliver desirable impact. Unfortunately, there were a few quantitative studies that represent to evaluate the existence of optical fiber networks in Indonesia.

Hence, this study aims to investigate whether the deployment of fiber optic cable of PRP can effectively narrow the connectivity gaps by region and improve economic activities. Connectivity is an internet activity that measured by using percentage of internet use and economic activity is measured by using Gross Regional Domestic Product (GRDP). Using a matching fixed effects (FE) staggered difference-in-difference (diff-in-diff) model on the panel data, we can explore the apparent causal effect of the non-random assigned policy in the case of PRP. This framework also allows us to explore the possible changes in internet use and economic activity when fiber optic cable of PRP arrives at various times.

This study appears that the existence of PRP infrastructure has a positive impact on the increasing percentage of individuals using the internet, but has no impact on GRDP. Note that the endogeneity bias may persist on the result as this study cannot fulfill the specific diff-in-diff assumptions. However, the result is robust after including several different sets of covariates to the specification. Finally, the findings and policy recommendations from this study are expected to contribute to the impact evaluation concerning the relationship among infrastructure development, internet disparities, and economic growth in order to improve the living standards and welfare.

2. Methodology

2.1. Data and variables

This study evaluates the impact of fiber optic internet backbones under PRP that gradually arrived during 2015-2020 in 46 districts in West, Central, and East Packages on the internet and economic activities. Note that the 46 districts represent the location of the PRP's landing points as presented on the *submarine cable map* website (https://www.submarinecablemap.com/). As the primary data source, this study utilizes the National Socio-Economic Survey (Susenas) March 2015-2020 of 46 districts conducted by Badan Pusat Statistik (BPS)-Statistics Indonesia (Table 1). This survey covered about 300,000 households in all the districts in Indonesia hence represented the data or information for each districts. In addition, this study also utilizes other source data from BPS-Statistics Indonesia, e.g. GRDP and number of population. The variables used in the model estimation, including dependent variables and explanatory variables that consist of PRP treatment and control variables, are presented in Table 2.

Table 1. List of Research Objects.

No	Area ID	Province	District	Landing Points	Packages
1	1408	Riau	Bengkalis	Bengkalis	West
2	1410		Kepulauan Meranti	Tebingtinggi Island	
3	1473		Dumai*	Dumai	
4	1507	Jambi	Tanjung Jabung Barat	Kuala Tungkal	
5	2101	Kepulauan Riau	Karimun	Karimun	
6	2103		Natuna	Natuna, Ranai	
7	2104		Lingga	Lingga	
8	2105		Kepulauan Anambas	Terempa	
9	2171		Batam*	Batam	
10	6172	Kalimantan Barat	Singkawang*	Singkawang	
11	7103	Sulawesi Utara	Kepulauan Sangihe	Tahuna	Central
12	7104		Kepulauan Talaud	Melonguane	
13	7108		Kep. Siau Tagulandang Biaro	Ondong Siau	
14	7171		Manado*	Manado	
15	7201	Sulawesi Tengah	Banggai Kepulauan	Salakan	
16	7202		Banggai	Luwuk	
17	7208		Parigi Moutong	Buranga	
18	7211		Banggai Laut	Banggai	
19	7402	Sulawesi Tenggara	Muna	Raha	
20	7412		Konawe Kepulauan	Wawonii	
21	7414		Buton Tengah	Lakudo	
22	7471		Kendari*	Kendari	
23	7472		Bau-bau*	Baubau	
24	8203	Maluku Utara	Kepulauan Sula	Sanana	
25	8205		Halmahera Utara	Tobelo	
26	8207		Pulau Morotai	Morotai	
27	8208		Pulau Taliabu	Taliabu	
28	8271		Ternate*	Ternate	
29	8272		Tidore Kepulauan*	Sofifi, Tidore	
30	5302	Nusa Tenggara Timur	Sumba Timur	Waingapu	East
31	5307		Alor	Kokar	
32	5314		Rote Ndao	Baa	
33	5320		Sabu Raijua	Seba	
34	5371		Kupang*	Kupang	
35	8101	Maluku	Maluku Tenggara Barat	Suemlaki	
36	8105		Kepulauan Aru	Kep. Aru	
37	8108		Maluku Barat Daya	Serwaru, Tiakur	
38	8172		Tual*	Tual	
39	9104	Papua Barat	Teluk Bintuni	Teluk	
40	9105		Manokwari	Manokwari	
41	9404	Papua	Nabire	Nabire	
42	9408		Kepulauan Yapen	Yapen	
43	9412		Mimika	Timika	
44	9414		Mappi	Kota Mappi	
45	9415		Asmat	Agats	
46	9427		Supiori	Supiori	

Note: *) municipality
Source: https://www.submarinecablemap.com/

The first variable of interest *inetuse* represents the percentage of individuals aged five and above who accessed the internet in the last three months. While the second variable of interest *lgrdp* represents the size of the economy for each district in natural logarithm form. The histograms in Figure A1 (see appendix) visually indicates that these two dependent variables are normally distributed.

Table 2. List of variables used in the model estimation.

Variable	Description	Source
Dependent:		
inetuse	percentage of internet users	Susenas, BPS-Statistics Indonesia
lgrdp	GRDP (ln)	publication and website of BPS-Statistics Indonesia
Explanatory:		
PRP (treatment)	PRP time-dummy variable: 1 – when PRP was ready for service in March of year *t* 0 – else	*submarine cable map* website
pov	percentage of poor population	Susenas, BPS-Statistics Indonesia
lpop	number of population (ln)	website of BPS-Statistics Indonesia
urban	percentage urban population	Susenas, BPS-Statistics Indonesia
mys	mean years of schooling	Susenas, BPS-Statistics Indonesia
phown	percentage individual own cellular phone	Susenas, BPS
eys	expected years of schooling	Susenas, BPS

The dummy variable of PRP indicates treatment effect caused by the fiber optic internet backbones under PRP. It takes value one since March in the closest year where PRP was completed or ready for service and all periods afterward; else zero. We use this strategy since our main source data was March Susenas, as such we consider the condition before and after the PRP was based on March. Assumption holds that the society could access the internet as soon as the PRP was finished. For example, PRP West package has been ready for service since February 2018, thus the dummy takes value one for 10 districts covered in 2018-2020; else zero. The 22 districts in the Central package value one for 2019-2020; else zero since the package has been completed since December 2018. Meanwhile, 17 districts in the East package value one for 2020 and zero for 2015-2019 as the package has been finished since October 2019. Unfortunately, due to the limitation on data availability, the data before treatment are greater—about 3 to 4 years— than the data after the project—about 1 to 3 years.

2.2. Estimation procedure

The several models would be employed to check the robustness of the regression coefficient estimates, especially for the dummy variable of PRP, that indicates the impact of PRP on the internet penetration (*inetuse*) as people are expected to have a higher digital engagement and on district's GDRP (*lgrdp*) as internet adoption could boost economic productivity. First, this study employs the naïve model as follows:

$$Y_{it} = \alpha + \tau PRP_{it} + \varepsilon_{it} \tag{1}$$

with *i* and *t* are subscripts for district and year, respectively and ε_{it} is error term. This model provides preliminary impact of PRP on both dependent variables through coefficient estimate τ, without considering impact from other auxiliary variables. Unfortunately, the estimates produced by this model are potentially biased due to heterogeneity in the panel data. In addition, the endogeneity issue also arise since the PRP appears to be partially non-exogenously designed as each package has different time, in terms of the project being completed and ready to use. This is a common condition as an infrastructure project, especially for national scale, first started from specific region [16]. As described in the previous sub-section that the first ready to service package is West package followed by the Central and East packages. By considering this condition, i.e. the PRP is staggered, to minimize the bias in baseline data this study only employs the districts that have PRP landing points in 2020, echoing the previous study [17]. Another strategy to reduce the policy endogeneity in panel studies was by employing region-fixed effects; however, this procedure somewhat will not remove the possibility of endogeneity issue from time-variant region characteristics [18].

Second, the full specification model used in this study is fixed effect (FE) staggered different-in-different (diff-in-diff) as follows:

$$Y_{it} = \alpha + \tau PRP_{it} + \beta X_{it} + \eta_i + t_t + \varepsilon_{it} \quad (2)$$

with X is a set of exogenous or explanatory variables with the impacts indicated by coefficient estimates β. Two fixed effects are incorporated in this model, i.e. district effect η that controls any time-invariant effects across districts and time effect t that controls any period shock. To produce unbiased estimated effect, FE staggered diff-in-diff model relies on parallel trend assumption that implies the parallel condition is appears for the time path of outcome variables in the absence of the project between treated and control group. The are several methods to test the parallel trend assumption, i.e. Granger-type causality tests, group-specific linear trends, and covariates balance test [19].

Third, the propensity score matching (PSM) in diff-in-diff model is utilized in this study to overcome parallel trend assumption by matching pre-treated outcomes of two groups before estimating the diff-in-diff model. This method could be performed for the non-randomized treatment to mitigate the time-constant unobserved variables in the two groups [20]. Finally, this study also performs heterogeneity analysis to check whether the PRP impact variation across the three packages.

3. Results and Analysis

3.1. Descriptive analysis

As previously mentioned, the PRP consists of districts that are grouped into three packages, i.e., Western, Central, and Eastern. The average changes of internet users over 2015-2020 by district and packages of PRP are presented in Figure 1. Generally, the districts covered in the PRP have experienced a positive impact on the internet users. From the figure, Manokwari (Papua Barat) appears to be the highest increase on the average change of percentage of internet users. Meanwhile, the lowest change was experienced by the district of Asmat (Papua). Interestingly, these two districts are in the same group, i.e., Eastern package. It indicates that over five years, the Eastern region has relatively more heterogeneous average change of internet users compared to the other packages.

On the contrary, the Western package appears to have a more homogenous average change of internet users compared to other packages. The districts in this region were more advanced in terms of the percentage of internet users, as the PRP Western package was operated earlier than the others. The figure shows that the district of Singkawang (Kalimantan Barat) has the highest average change in internet users, contrarily the district of Lingga (Kepulauan Riau) has the opposite condition. Meanwhile, in the Central package, the highest and lowest average changes in internet users are experienced by Buton Tengah (Sulawesi Utara) and Kepulauan Sula (Maluku Utara), respectively.

Figure 2 shows the percentage of internet users in 2020 as the last time reference of this study when all of the PRP packages were ready for operating. The districts of Batam (Kepulauan Riau), Kendari (Sulawesi Tenggara), and Kupang (Nusa Tenggara Timur) have the highest percentage of internet users in each package. Note that the two latter districts are the capital of the province. While the former is one of the districts with the most strategic location for international trade, industrial centre, and also tourism [21]. Moreover, in 2021, the Indonesian government has initiated the Special Economic Zones (SEZ) of Batam Aero Technic and Nongsa in Batam with the main activities in maintenance, repair, and overhaul (MRO) industry and also tourism and digital park to absorb the labors, attract the investment, and hence increase economic performance [22]. Nevertheless, related to the average change in figure 1, these three districts have a moderate average change in internet users over 2015-2020. Meanwhile, Lingga (Kepulauan Riau), Kepulauan Sula (Maluku Utara), and Asmat (Papua) are the districts with the lowest percentage of internet users in each package.

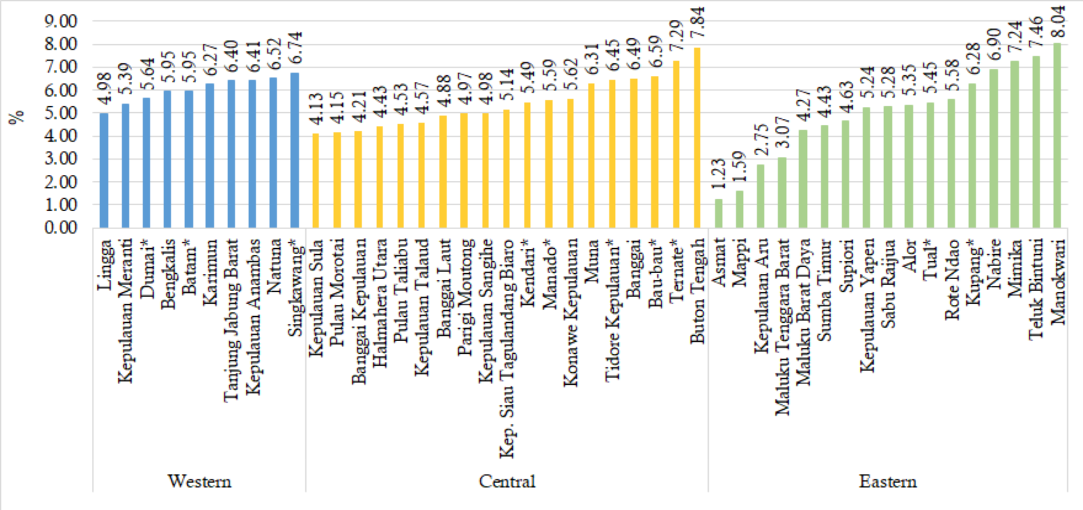


Figure 1. Average change of internet users by PRP package and district, 2015-2020 (%).
Note: *) municipality.
Source: Authors' calculation from Susenas March 2015-2020.

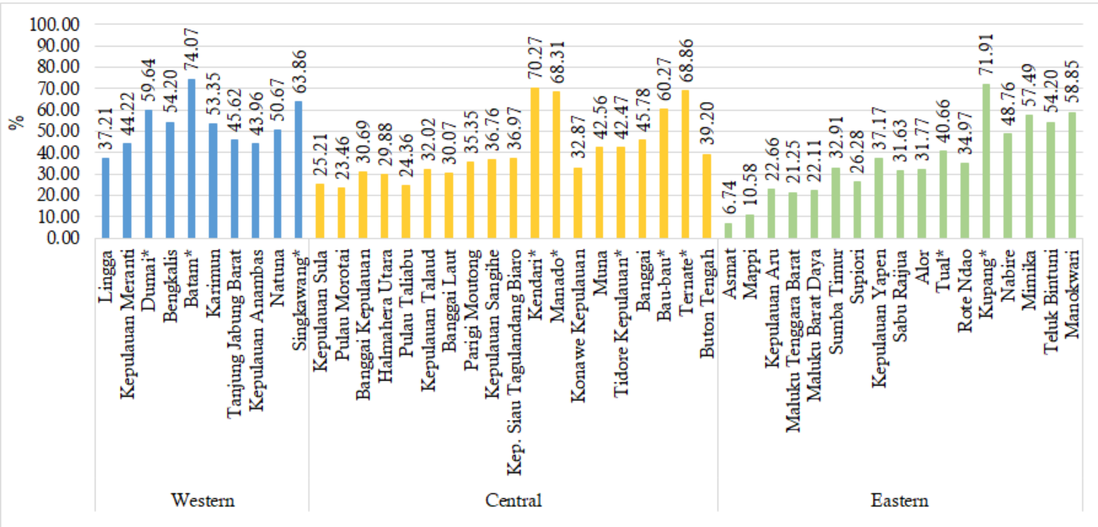


Figure 2. Internet users by PRP package and district, 2020 (%).
Note: *) municipality.
Source: Authors' calculation from Susenas March 2020.

Figure 3 describes the average economic growth by PRP packages and district during 2015-2020. While most of the districts in three packages have experienced a positive average economic growth, two districts in the Western package which are Kepulauan Anambas (Kepulauan Riau) and Bengkalis (Riau) likely to suffer from a negative average economic growth. Among the three packages, the Central region seems to have a higher level of economic growth with Banggai (Central Sulawesi) has the highest growth and Parigi Moutong (Central Sulawesi) has the lowest percentage. Meanwhile, the Central package appears to have a relatively homogenous range, with Mimika (Papua) holds the lowest average growth.

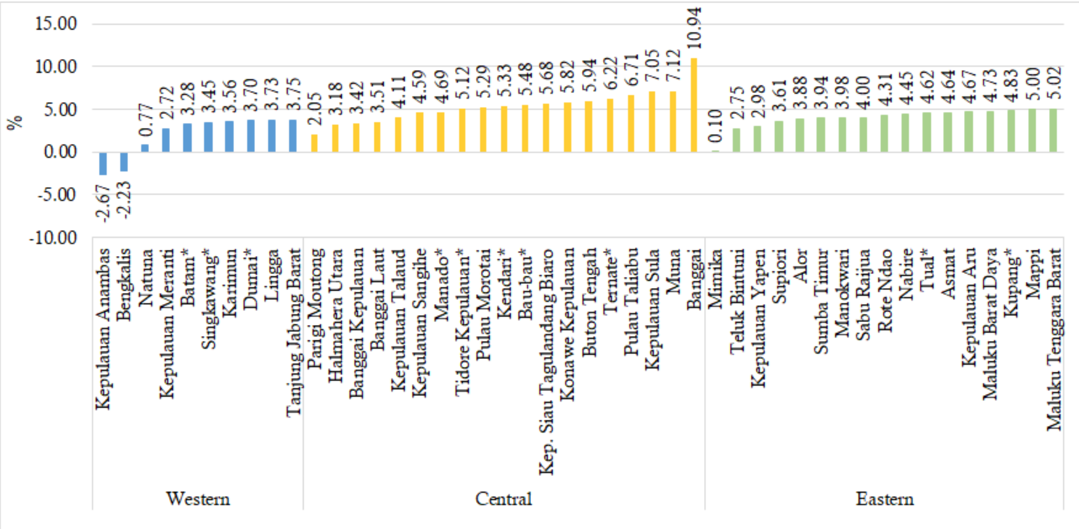


Figure 3. Average economic growth by PRP package and district, 2015-2020 (%).
Note: *) municipality
Source: publication and website of BPS-Statistics Indonesia 2015-2020.

Figure 4 presents the GRDP of all districts based on the PRP package year 2020. The figure shows that Batam (Kepulauan Riau), Manado (Sulawesi Utara), and Mimika (Papua) have the highest GRDP in each package. Batam is one of the leading districts in international trade and industry as its strategic location from Singapore and Malaysia. Manado is the second-largest economy in Sulawesi after Makassar. Meanwhile, the most substantial share contributing to Mimika's GRDP is the mining and excavation sector as the PT. Freeport Indonesia– a leading mining company – operated in this district. On the contrary, Lingga (Kepulauan Riau), Konawe Kepulauan (Sulawesi Tenggara), and Supiori (Papua) have the lowest GRDP in each package.

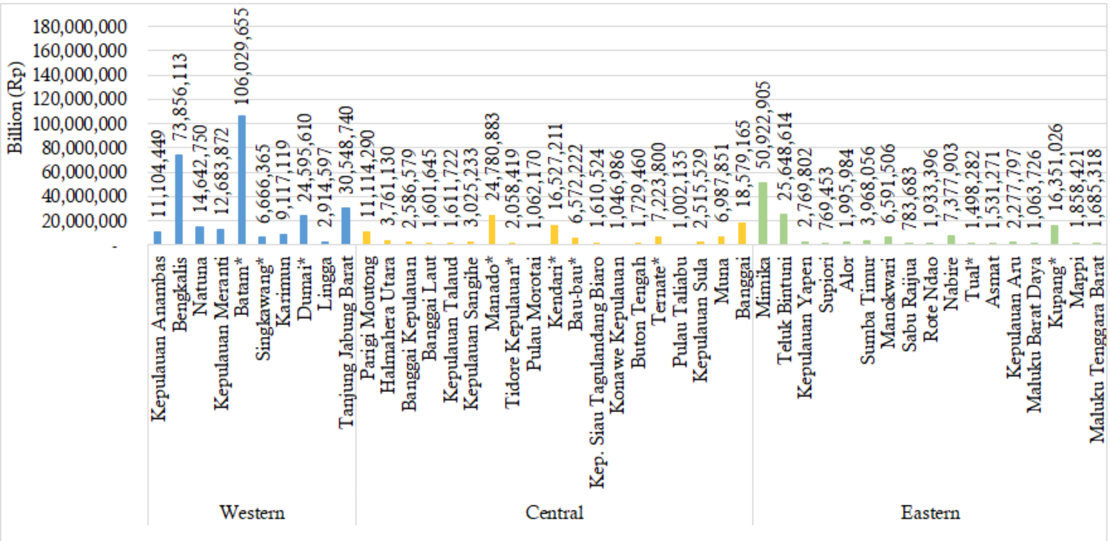


Figure 4. GRDP by PRP package and district, 2020 (billion Rp).
Note: *) municipality
Source: publication and website of BPS-Statistics Indonesia 2020.

3.2. FE staggered diff-in-diff model

The initial impact of PRP on the percentage of internet users and district's GRDP through naïve model is presented in Table 3. The result shows that the deployment of PRP has a positive significant impact on both variables of interest. The fiber optic internet backbones under PRP could increase the share of internet users up to 19,217% and boost the district's GRDP up to 44,10%. Since this naïve model produced a biased coefficient estimate, the impact of PRP need to be examined further using diff-in-diff models.

Table 3. Estimation result of naïve model.

Variable	*inetuse*	*lgrdp*
(1)	(2)	(3)
PRP (treatment)	19.217***	0.441***
	(1.994)	(0.169)
constant	21.188***	15.186***
	(1.057)	(0.092)
Observations	274	276
R^2	0.265	0.025

Notes: standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Source: Authors' calculation.

The multicollinearity check was performed before estimating diff-in-diff model. Table A2 (see appendix) shows pairwise correlation and Variance Inflation Factor (VIF). The table gives a view that all pairwise correlation coefficients between covariates are less than 0.8 and all the covariates have VIF below 10 indicating that the variables are free from multicollinearity issues. Hence, all the covariates are sufficient to be incorporated into the model.

Table 4 provides the estimation result of FE staggered diff-in-diff model both with and without control variables. To overcome arbitrary serial correlation over the period within groups, the standard errors are clustered at the district level [23]. By incorporating the control variables in the model, the coefficient estimates are expected to be more precise or have smaller standard errors [24]. The result shows that the existence of PRP significantly increases the percentage of internet users by 1.944% when the control variables are not included and by 1.707% when the control variables are included in the model indicates that the estimation of PRP impact is quite robust. On the contrary, the deployment of PRP has no significant impact on district's GRDP both when the control variables are included or not.

From six control variables, only the percentage of urban population that statistically significantly affects the percentage of internet users; 1 percent increase of urban population leads to the higher percentage of internet users by 0.42 percentage point. Meanwhile, there is no evidence to say that the PRP establishment significantly affected GRDP. This model also shows that all covariates are not statistically significant, except the percent of individuals who own cellular phones. A rise of 1 percent of individuals who own cellular phones would increase the GRDP by 0.40 percentage point.

Next, the parallel trend assumption checking is performed for both FE staggered diff-in-diff model with and without control variables using the Granger causality test and group-specific linear trend (Table A3 and A4 in appendix). The Granger causality performed by including the first lag of dummy variable PRP in the model. The results in Table A3 and A4 column (3) for model without control variables and column (6) for model with control variables show that all the coefficient estimates for the first lag of PRP are statistically insignificant indicate the anticipatory effect did not exist before PRP or parallel assumption holds. Otherwise, from the group-specific linear trend as provided in Table A3 and A4 column (4) for model without control variables and column (7) for model with control variables, the parallel trend assumption is not satisfied since the results are jointly significant, indicated by P-value of joint test less than 1%.

Table 4. Estimation result of FE staggered diff-in-diff model.

Variable	FE staggered diff-in-diff without control variables		FE staggered diff-in-diff with control variables	
	inetuse	*lgrdp*	*inetuse*	*lgrdp*
(1)	(2)	(3)	(4)	(5)
PRP (treatment)	1.944**	0.000	1.707*	0.001
	(0.942)	(0.016)	(0.891)	(0.014)
pov			-0.489	0.003
			(0.510)	(0.005)
lpop			4.713	-0.178
			(7.260)	(0.120)
urban			0.420**	-0.002
			(0.158)	(0.004)
mys			0.298	-0.005
			(2.148)	(0.045)
phown			0.028	0.004*
			(0.069)	(0.002)
eys			-0.589	-0.016
			(2.401)	(0.046)
District Fixed Effect	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes
Observations	274	276	280	238
R^2	0.893	0.703	0.792	0.256

Notes: clustered standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Source: Authors' calculation.

According to this mix results, the coefficient estimate of PRP impact may suffer from endogeneity bias, however, we are keen to provide the tools for robustness check. Thus, this study employs staggered diff-in-diff panel estimation by including one-by-one covariates to examine if the estimated treatment effects to outcome variables are relatively unchanged. Table A6 and A7 (see appendix) show the result of the PRP effect on percentage of internet users and GRDP. Both tables show that the estimated coefficients of *inetuse*, *lgrdp*, and all covariates are relatively unchanged in terms of number, direction, and significance as indication of robustness. Hence from the estimation results; the existence of PRP could provide a positive impact on internet penetration and brings no impact on GRDP.

3.3. Matching staggered diff-in-diff model

As the parallel trend assumption is violated, the combine method of matching using PSM and diff-in-diff model would be utilized as an alternative model. The inspection of balanced comparison between before and after matching is performed using the graph of standardized percentage bias of control variables (Figure A2 and A3 in appendix). The scattered and far from zero standardized percentage bias of these variables are depicted before and even after matching for the two outcomes, i.e. percentage of internet users (Figure A2) and GRDP (Figure A3), indicates that before and after matching are less balance as we have relatively small number of observations.

Table 5. Estimation result of matching diff-in-diff model.

Variable	*inetuse*	*lgrdp*
(1)	(2)	(3)
PRP (treatment)	1.870***	0.053***
	(0.000)	(0.000)
Control variables	Yes	Yes
constant	32.066***	15.390***
	(0.000)	(0.000)
Observations	8	8
R^2	1.000	1.000

Notes: both models also include control variables, district and time fixed effect. clustered standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Source: Authors' calculation.

Table 5 provides estimation of PRP impact on the percentage of internet users and GRDP from the matching diff-in-diff model. Both estimated impacts appear to be statistically significant at 1 percent significance level. However, these results may suffer from endogeneity biased since the graphs on standardized percentage bias did not provide substantially balanced comparison. Moreover, this specification only able to match eight observations, indicating that the matching diff-in-diff is not a reliable alternative. Since the matching diff-in-diff failed to meet its requirement, hence, the final model used in the analysis comes from FE staggered diff-in-diff model (Table 4).

3.4. Heterogeneity analysis

The impact of the existence of PRP infrastructure may vary across sections, therefore this study attempts to explore whether heterogeneity impacts the packages. Unfortunately, the small size of district observation causes matching staggered diff-in-diff for heterogeneity checking cannot be performed in this dataset. Furthermore, staggered diff-in-diff regressions are conducted to investigate heterogeneous impact across the packages. It is important to note that staggered diff-in-diff estimations possibility suffers from endogeneity bias. However, an initial view of the heterogeneity impact of PRP showed on the results.

Table 6. Heterogeneity effect of PRP using the staggered diff-in-diff model

Variable	*Inetuse*			*lgrdp*		
	West	Central	East	West	Central	East
(1)	(2)	(3)	(4)	(5)	(6)	(7)
PRP (treatment)	31.775***	21.022***	15.374**	0.192*	0.156***	0.261***
	(2.394)	(3.326)	(6.630)	(0.089)	(0.035)	(0.037)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	60	112	102	60	112	102
R-squared	0.977	0.939	0.846	0.556	0.910	0.758

Notes: both models also include control variables, district and time fixed effect. clustered standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Source: Authors' calculation.

The findings suggest that there were significant positive estimated impacts on internet use and GRDP in the packages. It indicates that there is a heterogeneity impact of PRP across regions. In particular, the largest effect of PRP infrastructure on internet penetration is on the West package is the largest then followed by Central and East. These findings confirm that the earlier the project is finished, the higher utilization of the network. In consequence, people in the West area are better at enjoying digital connectivity. Even so, it is important to highlight that these results do not hold exogeneity conditions. Moreover, the appearance of heterogeneity impact could be due to different baseline conditions. Aside from the well-existing infrastructure in West districts, one may notice that people living in western Indonesia are likely to have a better quality of life such as better digital skills, better education, better education, and readiness to use the new ICT compared to others. Even though the PRP infrastructure in the East package has been finished last year, the estimated effect on GRDP appears to be the highest. A plausible explanation is that less developed districts tend to experience economic growth leapfrog compared to more developed regions under the same category of infrastructure development [25].

Conclusion

The connectivity gap is a condition where internet access and infrastructure are unevenly distributed across regions. It is a common phenomenon in developing countries, including Indonesia. The telecommunication services providers believed that more finance need to build the infrastructure in the non-commercial and low profitable value of regions. In response, the government deploys PRP infrastructure or fiber optic internet backbones to provide infrastructure in order to make internet connection easier and faster thus narrow the digital gaps. Unfortunately, till PRP infrastructure was officially launched in the last year, there is a limited evaluation of the impact of PRP infrastructure on internet users and economic activities.

This study attempts to investigate the impact of the existence of PRP infrastructure on connectivity and economic activity by using FE staggered diff-in-diff model. Generally, from the descriptive analysis the districts covered in the PRP have experienced a positive impact on the connectivity and economic activity. This phenomenon supports the evidence from inferential analysis that the existence of PRP infrastructure has a positive and significant impact on increasing internet participation. In other hand, this phenomenon doesn't support for GRDP since the existence of PRP has no significant effect on enhancing GRDP. Though that estimation result is relatively robust, it is important to state that the results from the specification may suffer from endogeneity bias as this study fails to provide evidence to satisfy parallel trend assumptions. Hence, a PSM or matching FE staggered diff-in-diff was performed to mitigate the problem; unfortunately, the limited number of observations prevented us from having zero biases as there were only eight matching observations. Then, the estimation results from FE staggered diff-in-diff model are used as final estimation.

From the heterogeneity analysis, there is a heterogeneity impact of PRP on connectivity and economic activity depicted across packages. It also supported by the variation of percentage of internet users ad GRDP that visually appear across packages on descriptive analysis. These circumstances, perhaps, can be used as the preliminary information for the government to determine the priority areas for internet infrastructure development.

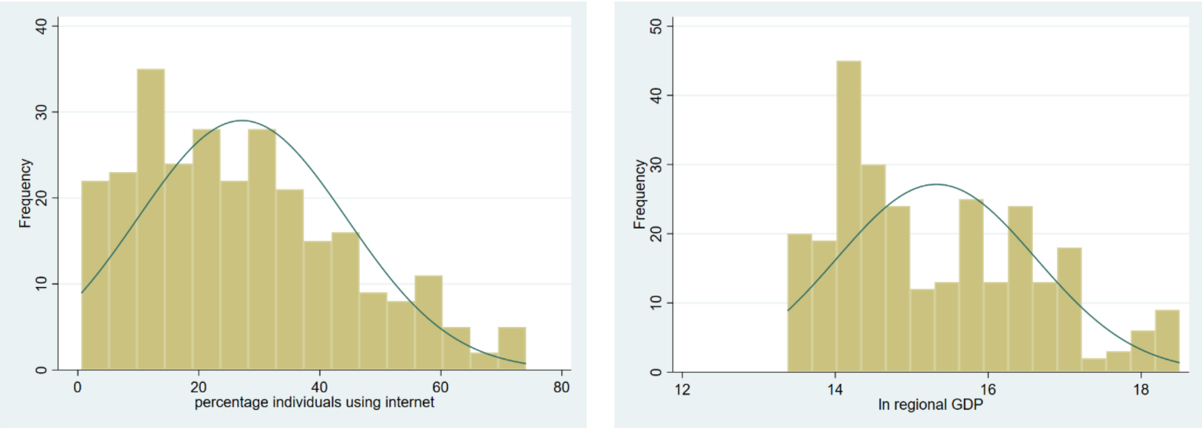
Referring to the result of the FE staggered diff-in-diff model, the establishment of PRP infrastructure has escalated the number of people using the internet; however, the effect is not yet able to enhance economic activities. Given that the project has just been launched, thus the second finding may be somewhat premature. The possible explanation why the mega-project fails to deliver an increasing on GDP is that people have limited knowledge and ICT skills, which restricts them to get the huge benefit of the internet. Consequently, they may prefer using the internet for entertainment purposes instead of seeking a job or earning a living.

In response to these findings, this study proposes several long-term policy recommendations. First, the policy should be designed with better collaboration and coordination between the policymaker and private sectors. Instead of being disincentives, the existence PRP infrastructure expected encourage commercial providers in exploiting the newly established infrastructure. While the center of attention tends to develop infrastructure policy, improvement quality such as coverage expansion and signal amplification should be taken to narrow the digital gaps. Second, while the urban population also plays a role in increasing internet users, the policy should focus on proliferating ICT facilities in rural areas. Third, since the percentage of phone use effect on GRDP, thus maintaining stability price of cellular phone, expanding education and digital literacy programs are also essential to improve economic activity as indicated by GRDP. Therefore, further studies can be carried out by exploring which specific GRDP sector experienced a higher growth.

Acknowledgments
The authors would like to thank BPS-Statistics Indonesia for support and facilities, Professor Blane D. Lewis – course convener of Quantitative Policy Impact Evaluation at the Australian National University – and Nurina Merdikawati for valuable lectures, helpful comments, and suggestion.

Appendices

Appendix A



Source: Authors' calculation.

Figure A1. Histogram of percentage of individual using the internet, and GRDP (ln).

Table A1. Summary statistics variables.

Variable	Before PRP (PRP=0)			After PRP (PRP=1)		
	Obs.	Mean	Std. dev.	Obs.	Mean	Std. dev.
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcomes						
% internet users (*inetuse*)	189	21.188	14.518	85	40.405	15.624
Ln GRDP (*lgrdp*)	191	15.186	1.277	85	15.627	1.310
Control Variables						
% poor population (*pov*)	190	17.417	10.230	85	12.412	8.461
Ln number of populations (*lpop*)	191	11.799	0.774	85	11.988	0.833
% urban population (*urban*)	189	36.457	27.859	85	45.085	28.283
Mean years of schooling (*mys*)	191	8.274	1.550	85	8.682	1.483
% individual own cellular phone (*phown*)	189	50.476	16.946	85	60.315	12.751
Expected years of schooling (*eys*)	191	12.532	1.423	85	13.023	1.139

Source: Authors' calculation.

Table A2. Pairwise correlation coefficients between the covariates and VIF.

Covariates	Outcome variable					
	pov	*lpop*	*urban*	*mys*	*phown*	*eys*
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pairwise correlation						
pov	1.0000					
lpop	-0.3869	1.0000				
urban	-0.4953	0.5604	1.0000			
mys	-0.4069	0.3894	0.6934	1.0000		
phown	-0.5547	0.5116	0.7486	0.6676	1.0000	
eys	-0.4116	0.3530	0.5292	0.6659	0.6252	1.0000
VIF	1.5040	1.5260	3.0230	2.6015	3.0423	2.0375

Source: Authors' calculation.

Table A3. Estimation result of FE staggered diff-in-diff and PSM staggered diff-in-diff model for percentage of internet users (*inetuse*).

Variable	FE staggered DID without covariates			FE staggered DID with covariates			Matching FE staggered DID
	Naïve model	Granger causality test	Group linear trend	Naïve model	Granger causality test	Group linear trend	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PRP	1.944**	2.584**	0.778	1.707*	3.172**	0.132	1.870***
	(0.942)	(1.238)	(0.759)	(0.891)	(1.188)	(0.695)	(0.000)
Lag PRP		0.710			1.070		
		(0.904)			(0.897)		
pov				-0.489	-0.554	-0.728	0.000
				(0.510)	(0.488)	(0.474)	(.)
lpop				4.713	56.575**	5.104	0.000
				(7.260)	(23.719)	(6.366)	(.)
urban				0.420**	0.403**	0.133	0.000
				(0.158)	(0.170)	(0.134)	(.)
mys				0.298	-0.762	5.201**	0.000
				(2.148)	(2.545)	(1.992)	(.)
phown				0.028	0.107	0.379***	0.000
				(0.069)	(0.091)	(0.073)	(.)
eys				-0.589	-0.402	-3.863**	0.000
				(2.401)	(2.026)	(1.911)	(.)
constant	14.787***	14.783***	-1912.205***	-44.488	-652.912**	-3499.435***	32.066***
	(0.577)	(0.502)	(94.550)	(78.981)	(278.747)	(711.664)	(0.000)
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
P-value joint test			0.000			0.000	
No. of observations	274	228	274	274	228	274	8
R^2	0.893	0.860	0.959	0.900	0.874	0.970	1.000

Notes: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Source: Authors' calculation.

Table A4. Estimation result of FE staggered diff-in-diff and PSM staggered diff-in-diff model for GRDP (*lgrdp*).

Variable	FE staggered DID without covariates			FE staggered DID with covariates			Matching FE staggered DID
	Initial model	Granger causality test	Group linear trend	Initial model	Granger causality test	Group linear trend	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PRP	0.000	-0.006	0.003	0.001	-0.004	0.000	0.053***
	(0.016)	(0.022)	(0.009)	(0.014)	(0.019)	(0.007)	(0.000)
F.prp		-0.011			-0.008		
		(0.009)			(0.007)		
pov				0.003	0.005	0.004	0.000
				(0.005)	(0.004)	(0.004)	(.)
lpop				-0.178	0.211	0.090**	0.000
				(0.120)	(0.208)	(0.042)	(.)
urban				-0.002	0.000	0.003	0.000
				(0.004)	(0.004)	(0.003)	(.)
mys				-0.005	-0.001	-0.031	0.000
				(0.045)	(0.047)	(0.033)	(.)
phown				0.004*	0.004	0.003	0.000
				(0.002)	(0.003)	(0.002)	(.)
eys				-0.016	-0.029	-0.004	0.000
				(0.046)	(0.053)	(0.033)	(.)
constant	15.198***	15.198***	6.175***	17.364***	12.787***	0.820	15.390***
	(0.009)	(0.008)	(1.048)	(1.836)	(2.366)	(8.122)	(0.000)
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
P-value joint test			0.000			0.000	
No. of observations	276	230	276	274	228	274	8
R^2	0.703	0.729	0.914	0.724	0.740	0.919	1.000

Notes: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.
Source: Authors' calculation.

Table A5. Estimation result of FE staggered diff-in-diff for covariate balance.

Variable	*pov*	*lpop*	*urban*	*mys*	*phown*	*eys*
(1)	(2)	(3)	(4)	(5)	(6)	(7)
PRP	0.185	-0.002	0.732	-0.007	-0.315	-0.063
	(0.203)	(0.009)	(0.516)	(0.027)	(0.789)	(0.045)
constant	16.789***	11.809***	36.426***	8.111***	48.248***	12.273***
	(0.155)	(0.003)	(0.277)	(0.021)	(0.478)	(0.033)
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	275	276	274	276	274	276
R^2	0.403	0.508	0.545	0.809	0.603	0.697

Notes: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.
Source: Authors’ calculation.

Table A6. Staggered diff-in-diff estimates of determinants of percentage of internet users (*inetuse)*

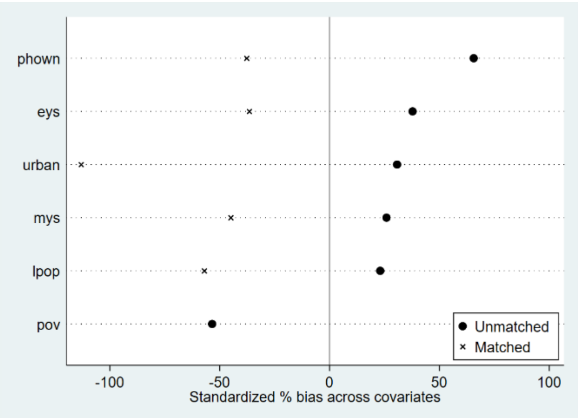
Variable	*inetuse*	*inetuse*	*inetuse*	*inetuse*	*inetuse*	*inetuse*
(1)	(2)	(3)	(4)	(5)	(6)	(7)
PRP	1.944**	2.028**	2.031**	1.740*	1.745*	1.707*
	(0.942)	(0.921)	(0.918)	(0.867)	(0.872)	(0.891)
pov		-0.471	-0.447	-0.528	-0.523	-0.489
		(0.534)	(0.505)	(0.504)	(0.509)	(0.510)
lpop			3.762	4.167	4.145	4.713
			(7.747)	(7.079)	(7.125)	(7.260)
urban				0.417**	0.414**	0.420**
				(0.159)	(0.158)	(0.158)
mys					0.381	0.298
					(2.154)	(2.148)
phown						0.028
						(0.069)
eys						-0.589
						(2.401)
constant	14.787***	22.700**	-22.132	-40.774	-43.545	-44.488
	(0.577)	(8.940)	(88.981)	(80.187)	(78.116)	(78.981)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	274	274	274	274	274	274
R^2	0.893	0.894	0.894	0.900	0.900	0.900

Notes: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.
Source: Authors’ calculation.

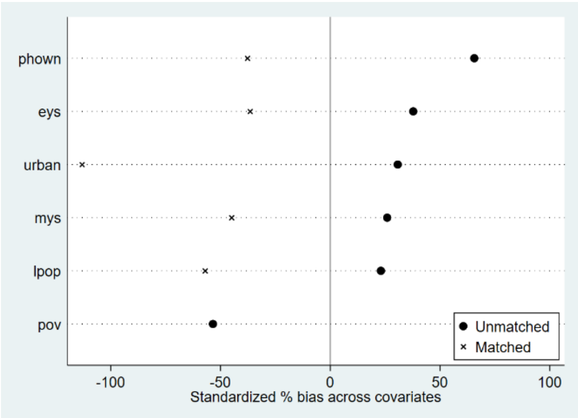
Table A7. Staggered diff-in-diff estimates of determinants of GRDP (*lgrdp)*

Variable	*lgrdp*	*lgrdp*	*lgrdp*	*lgrdp*	*lgrdp*	*lgrdp*
(1)	(2)	(3)	(4)	(5)	(6)	(7)
PRP	0.000	0.000	-0.000	0.002	0.002	0.001
	(0.016)	(0.016)	(0.017)	(0.016)	(0.016)	(0.014)
pov		0.001	-0.000	-0.000	-0.000	0.003
		(0.005)	(0.006)	(0.006)	(0.006)	(0.005)
lpop			-0.197	-0.199	-0.200	-0.178
			(0.176)	(0.175)	(0.176)	(0.120)
urban				-0.002	-0.002	-0.002
				(0.004)	(0.004)	(0.004)
mys					0.013	-0.005
					(0.043)	(0.045)
phown						0.004^{*}
						(0.002)
eys						-0.016
						(0.046)
constant	15.198^{***}	15.183^{***}	17.528^{***}	17.644^{***}	17.549^{***}	17.364^{***}
	(0.009)	(0.095)	(2.144)	(2.119)	(2.207)	(1.836)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	276	275	275	274	274	274
R^2	0.703	0.699	0.706	0.706	0.707	0.724

Notes: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.
Source: Authors' calculation.



Source: Authors' calculation.
Figure A2. Standardized percentage bias of covariate for percentage internet users.



Source: Authors' calculation.
Figure A3. Standardized percentage bias of covariate for GRDP (ln).

References

[1] W. Bank, "Individuals using the internet (% of population)-Indonesia," 2020. [Online]. Available: https://data.worldbank.org/indicator/IT.NET.USER.ZS?locations=ID%3E 11 October 2020.

[2] A. Barua, A. B. Whinston, and F. Yin, "Value and productivity in the Internet economy," *Computer (Long. Beach. Calif).*, vol. 33, no. 5, pp. 102–105, 2000.

[3] Google; and Temasek/Bain, "e-Conomy SEA 2019," *Google & Temasek/Bain*, 2016. [Online]. Available: https://www.blog.google/documents/47/SEA_Internet_Economy_Report_2019.pdf.

[4] McKinsey, "Unlocking Indonesia's digital opportunity," *McKinsey & Company*, 2016. [Online]. Available: https://www.mckinsey.com/~/media/McKinsey/Locations/Asia/Indonesia/Our Insights/Unlocking Indonesias digital opportunity/Unlocking_Indonesias_digital_opportunity.ashx. [Accessed: 13-Oct-2020].

[5] World Bank, "The digital economy in Southeast Asia: strengthening the foundations for future growth," *World Bank*, 2019. [Online]. Available: http://documents1.worldbank.org/curated/en/328941558708267736/pdf/The-Digital-Economy-in-Southeast-Asia-Strengthening-the-Foundations-for-Future-Growth.pdf. [Accessed: 13-Oct-2020].

[6] MCIT, "Ini skema kerja sama pemerintah badan usaha di palapa ring (The cooperation scheme between government and business entities in palapa ring)," *MCIT*, 2016. [Online]. Available: https://kominfo.go.id/content/detail/6997/ini-skema-kerja-sama-pemerintah-badan-usaha-di-palapa-ring/0/sorotan_media. [Accessed: 17-Oct-2020].

[7] N. Czernich, O. Falck, T. Kretschmer, and L. Woesmann, "Broadband infrastructure and economic growth," *Econ. J.*, vol. 121, no. 552, pp. 505–532, 2011.

[8] J. Kolko, "Broadband and local growth," *J. Urban Econ.*, vol. 71, pp. 100–113, 2012.

[9] M. B. Bulturbayevich and M. B. Jurayevich, "The impact of the digital economy on economic growth," *Int. J. Business, Law, Educ.*, vol. 1, no. 1, pp. 4–7, 2020.

[10] T. De Stefano, R. Kneller, and J. Timmis, "The (fuzzy) digital divide: the effect of broadband internet use on UK firm performance," 2014.

[11] J. Cariolle, "Telecommunication submarine-cable deployment and the digital divide in Sub-Saharan Africa," France, 2018.

[12] J. Hjort and J. Poulsen, "The arrival of fast internet and employment in Africa," *Am. Econ. Rev.*, vol. 109, no. 3, pp. 1032–1079, 2019.

[13] A. O'Connor, B. Anderson, A. Brower, and S. Lawrence, "Economic impacts of submarine fiber optic cables and broadband connectivity in Malaysia," NC, 2020.

[14] R. Adisti, "Human development impact of the implementation of broadband: a case study of the Indonesian palapa ring project," Delft University of Technology, Den Haag, 2017.

[15] C. Handoko and X. Zhang, "Case Study on The Palapa Ring Project: Prospects for Sub-National Competitiveness," 2021.

[16] D. McKenzie, "Strategies for evaluating the impact of big infrastructure projects: how can we tell if one big thing works?," *World Bank Blogs*, 2011. [Online]. Available: https://blogs.worldbank.org/impactevaluations/strategies-for-evaluating-the-impact-of-big-infrastructure-projects-how-can-we-tell-if-one-big-thing. [Accessed: 07-Nov-2020].

[17] J. Stevenson and J. Wolfers, "Bargaining in the shadow of the law: divorce laws and family distress," *Q. J. Econ.*, vol. 121, no. 1, pp. 267–288, 2006.

[18] T. Besley and A. Case, "Unnatural Experiments? Estimating the Incidence of Endogenous Policies," *Econ. J.*, vol. 110, no. 467, pp. 672–694, 2000.

[19] C. Wing, K. Simon, and R. A. Bello-Gomez, "Designing difference in difference studies: best practices for public health policy research," *Annu. Rev. Public Health*, vol. 39, pp. 453–469, 2018.

[20] P. Gertler, S. Martinez, P. Premand, L. B. Rawlings, and C. M. Vermeersch, *Impact evaluation in practice*. Washington DC: The World Bank, 2011.

[21] Zanuardi and Kudariyanto, "Persepsi Masyarakat dalam Ketersediaan Penyediaan Infrastruktur terhadap Tingginya Laju Pertumbuhan Penduduk Kota Batam [Community Perceptions in the Availability of Infrastructure Provision on the High Population Growth Rate of Batam City]," *J. Sosek Pekerj. Umum*, vol. 2, no. 3, pp. 151–157, 2010.
[22] Indonesia National Council of SEZ, "Pemerintah Tetapkan dua KEK Baru di Batam [Indonesian Government Initiates Two New SEZs in Batam]," 2021. [Online]. Available: https://kek.go.id/berita/2021/06/Pemerintah-Tetapkan-Dua-KEK-Baru-di-Batam-298.
[23] B. Lewis, "Does local government proliferation improve public service delivery? Evidence from Indonesia," *J. Urban Public Aff.*, 2017.
[24] B. Lewis, "IDEC8026 Quantitative policy impact evaluation," *Lect. 31 August, slide 18, Crawford Sch. Public Policy, Aust. Natl. Univ.*, 2020.
[25] J. Lin, Z. Yu, Y. Wei, and M. Wang, "Internet access, spillover and regional development in China," *Sustainability*, vol. 9, no. 946, pp. 1–18, 2017.