Social Fractionalization and Economic Development:

Evidence from the Korean *New Village Movement*[*]

BooKang Seol[†]

(Job Market Paper)

February 9, 2023

Click here for the latest version

Abstract

Social fractionalization along dimensions like ethnicity or class can stunt economic development. This paper investigates how fractionalization affects a group's ability to respond to economic shocks by studying rural South Korea in the 1970s. Social groups in rural Korea were defined by one dominant characteristic: extended kin network identified by family name. Some villages displayed high homogeneity, with as many as 90% of households sharing a family name. This unique social context offers a reliable measure of social fractionalization that is otherwise complex and difficult to measure. I combine this cross-sectional variation with the time variation in market access created by the construction of a new bridge, the *Namhae* bridge, in 1973. I find that homogeneous villages displayed higher agricultural investments and productivity growth than heterogeneous villages following the bridge construction. Homogeneous villages better exploited opportunities created by the bridge by providing complementary local public goods more effectively than heterogeneous villages. Finally, heterogeneous villages did catch up, consistent with spillovers or demonstration effects.

[*]I am grateful to Miguel Urquiola, Suresh Naidu, and Eric Verhoogen for their guidance and support throughout the Ph.D. program at Columbia. I also thank Jie Bai, Michael Best, Sandra Black, Do-Hyun Han, Sokbae Lee, Cristian Pop-Eleches, Matthias Schündeln, Rodrigo Soares, Jack Willis, and Hyunjoo Yang for their valuable comments. I benefited from discussions with Omar Ahsan, Sujoy Bhattacharyya, Naman Garg, Florian Grosset, Vince Jeong, and Krzysztof Zaremba. Special thanks to Mr. Euiyeon Jeong and Ms. Heemyung Ahn for generously providing me with the *Namhae* data they had collected. Sanghwa Ahn, Chris Babcock, Yuqi Cheng, James Hu, Sujeong Kim, Seoyeon Kim, Hyungcheol Kim, Taein Kim, Byeongjin Min, Bukyung Min, ChaeYeon Oh, Jiyoung Park, Siyeon Song, and Abby Zhang provided excellent research assistance at various stages of the project. I would like to thank Miguel Urquiola, The Program for Economic Research (PER), The Center for Development Economics and Policy (CDEP), and the Colloquium Research Fund at Columbia for their financial support.

[†]Department of Economics, Columbia University. bookang.seol@columbia.edu

1 Introduction

Social fractionalization refers to divisions within society along dimensions like race, ethnicity, religion, language, or culture. A high level of fractionalization can negatively impact economic development (Alesina et al. 2003, Hjort 2014, Burgess et al. 2015). Yet, less is known about how social fractionalization impacts a group's ability to respond to an economic shock and drive productivity growth. This paper investigates this question by studying South Korea in the 1970s when the country experienced a 60% growth in agricultural output per household over ten years.

Rural Korea in the 1970s is an ideal environment for studying this question for three reasons. First, it offers a reliable measure of social fractionalization at the village level. Typically, measuring fractionalization is difficult because fractionalization tends to occur along many dimensions simultaneously to define a single variable to proxy social structure (Bossert et al. 2011, Esteban and Ray 2011, Esteban et al. 2012). In contrast, the fractionalization of Korean villages can be proxied by the geographic concentration of family clans that shared the same parental lineage. This is because Korean villages were defined by one dominant characteristic: extended kin network. Other social divisions prevalent in developing countries, such as ethnicity and language, did not exist in Korea. Clans were a legacy of a pre-modern, class-based society, representing social fractionalization that persisted in rural villages into the 20th century. Some villages displayed high clan homogeneity, with as many as 90% of households sharing a family name. A major reason for the high concentration of clans was the change in inheritance law in the 17th century, which excluded daughters from being inheritors. Daughters were then incentivized to leave their native villages to live with their husbands, while sons often lived close to inherited assets, such as farmland or a house (Yang 2019).

Second, economic outcomes differed meaningfully across villages because each village had significant decision-making power. The Korean government delegated authority for public good provision to local communities in the 1970s. This decentralized approach to rural development allowed villages to adjust the provision of public goods according to their priorities. Participating communities formed village councils, and elected local leaders sought a consensus on project priorities and led the projects. It is known as the *New Village Movement* in Korea.

Lastly, an economic shock created a natural experiment across villages: an uneven expansion of road networks. I focus on the construction of the *Namhae* bridge that connects Korea's fourth largest island, *Namhae* island, with the *Namhae* Expressway that runs through the major metropolitan areas in the Southern states. The bridge and the expressway together significantly reduced the travel time from the island to the nearby cities and provided stable access to road networks, regardless of sea conditions.[1] Both the bridge and the expressway were completed in 1973.

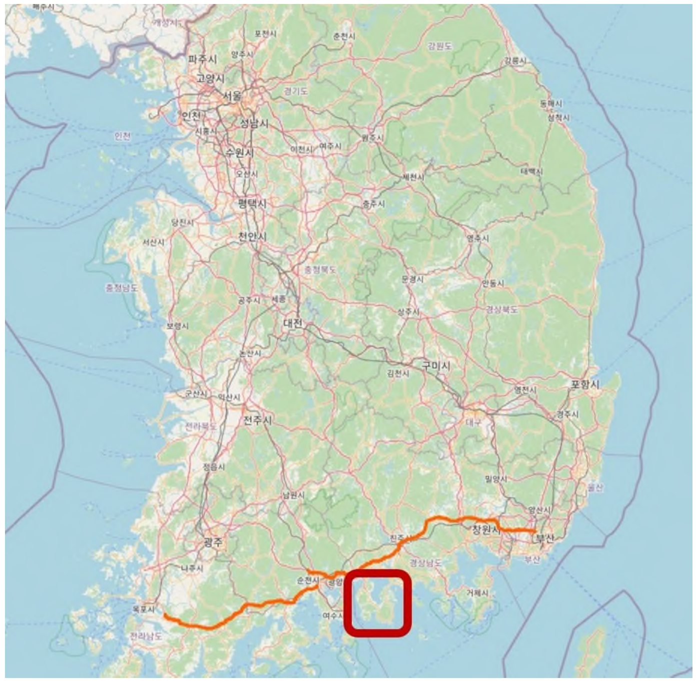
These three factors together create the variations necessary to identify the causal impact of fractionalization on economic outcomes. The social fractionalization based on family clans created a cross-sectional variation. The expansion of road networks created time variation in market acccess. Combining both sources of variation allows me to observe how villages responded differently to the market access shock based on their underlying social fractionalization.

Figure 1a shows a map of South Korea with the *Namhae* Expressway highlighted in orange and *Namhae* county in a red rectangle. Figures 1b shows *Namhae* county with its two islands and eight townships.[2] I use the smaller island without access to the bridge as the control in my main specification. In 1973, the connected island had 104,359 residents (18,474 households) in 169 villages. The unconnected, smaller island had 20,315 residents (3,233 households) in 26 villages. Figure 1c shows a view of the *Namhae* bridge in 1973.

[1]The construction of the *Namhae* bridge is considered a historic moment for the residents of the island. A Korean historian at Kookmin University, Kim Young-Mi, who was born and raised in *Namhae* county, described that traveling from *Busan*, the second largest city, to the *Namhae* island was a "scary" adventure before the bridge was built (https://youtu.be/ZEUeRi_F_oI). The trip would take her a whole day to travel to the nearest village across the island, where she had to sleep at a hostel to catch the next morning's ferry. She described being able to travel to *Busan* in a few hours by bus as "a historic improvement."

[2]A township subdivides into villages. A typical township has about 25 villages under its jurisdiction.

Figure 1: The *Namhae* Expressway, *Namhae* Bridge, and *Namhae* county



(a) The *Namhae* Expressway is highlighted in orange, and *Namhae* county is highlighted with the red rectangle.

(b) *Namhae* county and its eight townships

(c) The *Namhae* bridge in 1973

To measure economic development at the village level, I constructed a unique data set covering 195 villages in *Namhae* county between 1969 and 1984. I digitized over 336 volumes of historical records that document economic activities at the village level. To my knowledge, this is the first paper that uses granular data to study Korea's successful rural development in the 1970s.

I focus on three aspects of rural development. First, I examine agricultural productivity growth using two measures: rice yield and rice output per agricultural household. Rice is the main staple crop in Korea, making up 40% of the total agricultural production by weight in 1974 in *Namhae* county (Namhae County Office 1974).

Second, I examine each village's adoption of agricultural technology—specifically, mechanical tillers. Mechanical tillers were becoming widely available in rural villages in the 1970s, and a large body of anecdotal evidence supports their positive impact on agricultural productivity. A tiller has many uses, ranging from stirring and pulverizing the soil before planting, to pumping water, to transporting materials. To this day, tillers are widely used in rural areas as multi-purpose vehicles.

Lastly, I examine public good provision by village. Every year, the government evaluated villages' performance in providing public goods using a three-tier grading scale (A, B, and C).[3] Figure A.1 shows an example page of the evaluation sheet. The evaluation was thorough: a typical evaluation had more than 50 categories, ranging from investments in local infrastructure to the frequency of village meetings. Villages also took them seriously, as future government support depended on their letter grades. Besides the village grades, I also look at the total length of village roads constructed, the number of village project participants, reservoir capacity, and the amount of home-produced fertilizers.

My overall finding is that less fractionalized villages saw better economic outcomes, largely due to faster technology adoption and better collective action. Homogeneous villages that gained access to the new bridge achieved higher agricultural productivity growth than both heterogeneous villages and other homogeneous villages that did not have access to the bridge. This higher growth was driven by faster technology adoption and the use of

[3]The original classifications were *Jarib* (independent), *Jajo* (self-help), and *Gicho* (basic) villages. Following Yang (2019), I replaced these labels with A, B, and C respectively.

fertilizers.

Additionally, I find that homogeneous villages invested more in complementary public goods to exploit opportunities created by the new bridge than heterogeneous villages. When the bridge was initially constructed, village and farm roads were neither paved nor wide enough to accommodate modern vehicles like tillers. Figure A.2 shows a picture of a typical rural village in 1973 and a picture of a tiller in the same year. When given the authority to allocate resources, the leaders from homogeneous villages chose to invest in road infrastructure. They were also able to garner more support from the villagers, as shown in the number of project participants per household. Consequently, the homogeneous villages spent more years with the "A" grade status than heterogeneous villages. These findings are consistent with the fact that homogeneous villages held more village meetings, and more people showed up to them, albeit primarily men. Finally, I find that heterogeneous villages did catch up, consistent with spillovers or demonstration effects.

The results are robust to various geographic, pre-bridge village characteristics and village-level controls. The baseline set of covariates includes controls for geographic and pre-bridge village characteristics such as population, the share of agricultural households, the age of village leaders, the share of modified roofs, the number of schools, car accessibility, access to water, total land area, rice output, and a dummy for whether a village has a stream and a dummy for villages with mountains. I also present results with township-time fixed effects for additional flexibility.

This paper contributes to a few strands of literature. First, it highlights a specific way in which social fractionalization may negatively affect economic development. This insight adds color to the growing body of research on social fractionalization and economic development (Hjort 2014, Burgess et al. 2015, Moscona et al. 2020).

This paper also offers a nuanced perspective on the effect of transport infrastructure in rural areas, where the literature has shown inconsistent results. While Asher and Novosad (2020) saw that rural road construction did not necessarily improve agricultural outcomes in India, Shamdasani (2021) found the opposite. This paper shows how local context contributes to the differential outcomes: the effect of transportation on rural development depends in part on group's ability to exploit new opportunities. Local context is particularly salient for rural transport infrastructure projects because of the well-known “last mile” problem. This insight is in line with Gebresilasse (2021), which emphasizes the importance of complementary policies and investments, to fully reap the benefits of investments in rural areas.

Additionally, this paper adds a perspective to the mixed evidence around the impact of delegating authority on public goods and services provision. On the one hand, bottom-up planning could facilitate better use of local information and higher ownership. Alatas et al. (2012) find that allowing the communities to select program beneficiaries resulted in higher satisfaction. On the other hand, such practice creates risk for elite capture, per research by Heß et al. (2021). In the case of 1970s Korea, delegating authority to villages enabled local communities to take advantage of new economic opportunities and significantly improve agricultural productivity.

Lastly, this paper is one of the first attempts to demonstrate the causal effects of Korea’s *New Village Movement*, a phenomenon that Korean academics from sociology and history to social studies are actively investigating to this day.

2 Context and data

2.1 Background and natural experiment

In the 1970s, the Korean economy was undergoing a structural transformation; the economic activities shifted from agricultural to manufacturing. The value added of Agriculture, Forestry, and Fisheries as a share of GDP halved from 26.5% in 1970 to 14.3% in 1980 (Korea National Statistics Office 2022).

Yet, such a description masks the fact that Korea’s structural transformation accompanied a 60% growth in agricultural productivity. Figure A.3 plots rice output per agricultural household in green and rice yield in grey between 1965 and 1979. I normalized the values from 1970 to 100. In less than a decade, both proxies of agricultural productivity jumped by 40% to 60%.

This paper investigates the factors that drove this agricultural growth. In particular, I focus on how pre-existing social fractionalization based on lineage interacted with economic policies to create this growth.

I exploit two sources of variation. The first is cross-sectional variation that stems from family clan homogeneity. The formation of clan villages has been a long-standing question in Korean history. Historians and sociologists attribute the reason to a change in the inheritance law in the 17th century (Yang 2019). Before the reform, both sons and daughters had equal inheritance rights. When the inherited assets were immovable, daughters were equally likely to remain in the same village after marriage. This would increase clan heterogeneity in the village, as daughters, once married, belonged to the husband's family name. However, after the reform, it became customary for daughters to leave their natal villages after marriage. This pattern increased the concentration of clans over a long period. By 1930, about half of all villages in Korea were considered "clan villages," most of which were first formed more than 300 years ago (The Academy of Korean Studies 1991). This resembles tribal cultures found in other countries, though the force that bonded the clan groups in Korea is solely based on kinship. Because the settlement pattern is likely related to geographical characteristics, such as soil quality or access to water, I control for these variables in my main specifications.

The second source of variation comes from the two concurrent economic policies in the 70s: the rapid expansion of transport infrastructure networks and decentralization of public good provision. The government pursued ambitious national transport infrastructure projects during this decade. In 1970, the first large-scale highway (the *Gyeongbu* Expressway) that spanned 428 km between the two largest cities, Seoul and Busan, was completed. By the end of 1973, two additional major highway networks were completed, adding 533 km to the road networks (the *Honam* Expressway (260.7 km) and the *Namhae* Expressway (273.17 km)). The *Namhae* bridge, which connects *Namhae* island with the *Namhae* Expressway, was also completed in 1973. I focus on the construction of the *Namhae* bridge, which created time variation in market access.

The rapid expansion of transport infrastructure overlapped with the launch of a rural development campaign called the New Village Movement. It was a nationwide, community-driven development program that provided government transfers conditional on using them for public good provision. The program decentralized public good provision by delegating the authority for fund allocation to village leaders. Villagers carried out all phases of the projects—from planning and implementing to monitoring them. They matched government transfers with voluntary labor and material contributions.

In 1970, the campaign started by distributing cement and rebar for small-scale infrastructure projects in rural villages. Almost every 36,000 villages participated in the campaign by the decade's end. Public documents from the 70s often described that it "spread like wildfire" (The Ministry of Home Affairs 1978). Between 1970 and 1979, 43,506 km of village roads and 61,201 km of farm roads were constructed and paved. 35,608 village centers were built, which means that almost every village rebuilt its village center (The Ministry of Home Affairs 1980). Thatched roof, once the epitome of rural villages, was a thing of the past by 1980. The movement's momentum continued until 1979, but it lost its steam with the assassination of President Park Chung-hee.

The most notable achievement of rural development in Korea is that an agricultural household's income exceeded that of an urban household between 1974 and 1978. Figure A.4 shows the income gap between urban and agricultural households. The agricultural income continued to grow on par with urban income until 1984.

There remain controversies over the level of discretion the villages had at the time, but the *New Village Movement* was a grassroots movement at its heart. Village leaders were responsible for reaching a consensus on project priorities, encouraging villagers to participate, and budgeting the projects. Their suggestions were frequently rejected by villagers (Han 2010). A few unique institutional setups separate the New Village Movement from the more contemporary community-driven development practices, which could have contributed to its success. However, what is relevant for this paper is that the New Village Movement provided flexibility for the villagers to decide their priorities. Figure A.5 shows the evolution of the project budgets in Namhae county in the 1970s. Villager's contributions made up most of the budget. This would have been difficult to achieve if they did not enjoy substantial direction over the choice of projects.

The combination of clan homogeneity, national transportation projects, and decentralized public good provision together creates a unique environment to study Korea's rapid agricultural productivity growth. Expanding road networks created variation in market access, while the decentralization policy provided resources and legitimacy to villagers to self-organize and address their most pressing needs. However, not every village was well-positioned to take advantage of the increased market access and the decentralization policy. Some had strong leadership based on kinship, while others did not. This paper aims to quantify clan homogeneity's relevance in explaining Korea's agricultural growth in the 70s.

2.2 Data sources and digitization

Despite the significance of Korea's rural development in its history, only a few village-level records from the 70s remain today. It is difficult to make a statistically meaningful inference with publicly available data. As such, I searched for historical records that documented economic development at the village level. The challenge was to find the records that were consistent over time and remained at a reasonable scope to make an empirically meaningful sample size.

I digitized and combined data from two new sources from Namhae county: data related to village development projects from *History of Our Village* and village-level output data from *Status Reports of Townships*. To expedite digitizing 336 volumes (around 33,000 pages) of historical archival data and to work within financial constraints, I used an off-the-shelf OCR engine to automate the process. Recent advances in deep learning have significantly improved the accuracy of automated digitization and automation has recently gained more attention in social sciences (Shen et al. 2020). The first draft of digitization was then manually reviewed by a company in India for errors. I also hired RAs to further check for errors by comparing the recorded sub-totals with actual sums in each column. An example page of the OCR-ed page is shown in Figure A.6. The search and digitization efforts produced Korea's first village-level panel data between 1969 and 1984.

History of Our Village. I get data related to village projects from *History of Our Village*. Figure A.7 shows the cover and an example page. In 1978, the county mayor distributed a history book template to village leaders and asked them to fill it out. Some villages kept updating them until the 2000s. Of the 195 villages in Namhae county, 122 villages still had them archived. The history books contain a detailed record of the village projects since 1971. It documented each project's total budget, the amount of villagers' voluntary contributions, the number of participants, and other project details, such as the length and width of the newly constructed roads. Village status was also recorded: population, village's joint assets, the number of agricultural machinery, cultivated areas, and the number of TVs, etc. To my knowledge, this is the only surviving document with the list of village projects recorded at the village level.

The validity of data from the history books can be questioned, as they were not audited. I provide a few pieces of evidence that give me confidence in their validity. First, though the village leaders wrote the history books, it was an initiative of the county mayor. In the months following the distribution of the template, he visited every village, signed the books, and left short messages. Second, the history books were first filled out in 1978, when most project-related documents still existed. The village leaders were required by law to maintain project records for ten years. Lastly, township offices used to keep track of village projects using project "report cards." I found that the report card entries matched the history books' entries. Figure A.8 shows a report card for a farm road project in 1975 from the archives and a page from the corresponding village's history book. The project details in this report card match exactly with the history book's documentation. Though it validates only one project, it provides a clue for where the history books' project records originated.

Status Reports of Townships. The main village-level output variables are drawn from the *Status Reports of Townships*. On average, 25 villages formed a township. Township offices were required to report annually to the county office on administrative statistics of the villages under their jurisdictions. The statistics included information about demographics (population by age group or industry sector), public finance (tax revenues and expenditures), agricultural inputs and outputs (use of fertilizers, machinery, and production of rice and cereals), and other socioeconomic statistics such as education (the number of schools, teachers, and students by class). These reports became the basis for the Statistical Yearbook, an annual county publication of administrative statistics. As such, data were collected through standardized forms. The advantage of using these reports is that the statistics are reported at the village level. The status reports were missing for the years between 1970 and 1974, except for the two townships, which together comprised 25% of the villages in the county (49 villages out of 195 villages).

For my primary analysis, I use a balanced panel of 195 villages for 1969, 1975-1979, 1981, 1982, and 1984. When the project-related data from History of Our Village is used, the number of villages is reduced to 122 villages.

Other Complementary Data. I use the village-level data from the following two additional sources: Family clan data from the 1930 Census and initial village evaluation results from the *New Village Comprehensive Survey* published in 1972. According to the 1930 Census, 40 out of 195 villages in Namhae county had more than 50% of the households belonging to a single clan. I provide more detail about the family clan data in the following section.

Lastly, the *New Village Comprehensive Survey*, published in 1972, contains the first evaluation results of all villages in Korea. In addition to the letter grades that the villages received, it includes village characteristics such as car accessibility and a village's access to water and electricity.

Table 1 reports descriptive statistics for village projects and village-level characteristics. The top panel shows the summary statistics of village project data from the *History of Our Village*. On average, a village completed 4.5 projects per year and about 36.8 projects between 1970 and 1979. The share of villages that receive a grade of A steadily increased from 12% in 1971, 25% in 1975, and 100% in 1979. The share of the budget spent on village infrastructure also exhibited a similar pattern, increasing from 52% in 1971 to 72% in 1979.

The bottom panel shows the summary statistics of village-level data from the *Status Reports of Townships* and the *1930 Census* on family names. As in other parts of rural Korea, *Namhae* county was experiencing out-migration. Yet, the number of households fell slower, implying that the size of households reduced. The cultivated land area stayed constant, while rice output increased by 50%. The number of mechanized tillers started to be tracked in a small number of villages in 1975, with more widespread tracking beginning in 1976. By 1979, the average number of tillers per village was less than six. Given that a typical village had 100 households, fewer than 6% of households owned a tiller even by the end of the 70s.

Table 1: Summary Statistics of the Two Main Data Sources

Panel A	(1)	(2)	(3)	(4)	(5)
Village Projects	Pooled-Years			By Year	
	Mean	SD	1971	1975	1979
Num. of Projects	36.8	23.4	4.02	4.46	3.42
Avg. Budget per Project	\$1,448	\$2,146	\$546	\$1,182	\$2,737
Share of Budget on Infrastructure	56%	33%	52%	53%	72%
% of Villages with Grade A			12%	25%	100%
Observations (# *Villages*)	5,278 (120)		528 (112)	528 (112)	429 (102)

Panel B	(1)	(2)	(3)	(4)	(5)
Village Characteristics	Pooled-Years			By Year	
	Mean	SD	1969	1975	1979
Population	591.6	313.0	689.5	561.2	565.7
Household	111.5	59.5	114.7	113.7	109.1
Agri. Household	92.2	37.6	92.3	94.0	84.1
Cultivated Areas (ha)	25.6	12.4	25.1	25.3	25.5
Rice Production (tons)	99.7	53.5	76.1	89.1	113.9
Tillers	3.68	3.06	N/A	1.93	5.61
Observations (# *Villages*)	1,370 (196)		(195)	(196)	(196)

2.3 Measuring Social Fractionalization

This study requires a measure of village fractionalization that reflects its ability to coordinate. The challenge is that fractionalization is an abstract concept and cannot be observed directly. I proxy for fractionalization using the Herfindahl index of family clan concentration. It is a suitable proxy for two reasons. First, the defining characteristic of rural villages in Korea was the existence of family clans. When a village had a dominant clan, a clan member was often elected as the village leader and led village-level decisions. This is reflected in the family names of past village leaders; clan-dominant villages often have most past leaders share the same family name dating back to the Japanese colonial period. Second, Esteban et al. (2012) provide a theoretical motivation for using the Herfindahl index. They argue that it captures the probability that two randomly chosen individuals belong to different clans.

The family clan data comes from the 1930 population census conducted by the Japanese Colonial Government.[4] Figure A.9 shows an example page of the Census. It recorded the

[4]Professor Hyunjoo Yang at Sogang University provided me with this data. He digitized the Census and

number of households with a given family name and ancestral origin that comprised more than 10% of the total households in a village. Both pieces of information are relevant for defining clan in the Korean context; for example, people with last name Park from the city of Mil-Yang are considered to be of a different clan than people with last name Park from the county of Muan.

Let s_{ic} denote the share of households that belong to a family clan c in the village i. Then, the Herfindahl-Hirschman index of village i, H_i, is calculated as:

$$H_i = \sum_{c=1}^{N} s_{ic}^2$$

I complement the Herfindahl index with two additional measures: $TOPSHARE_i$ and $CLANSHARE_i$. TOPSHARE is the household share of the largest clan in a village. CLANSHARE is the sum of household shares of the clans that comprise more than 10% of the total households in a village. Table 2 shows summary statistics of the three measures. I also provide the spatial distribution of the Herfindahl index of clan density in Figure 2.

Table 2: Three Measures of Social Fractionalization

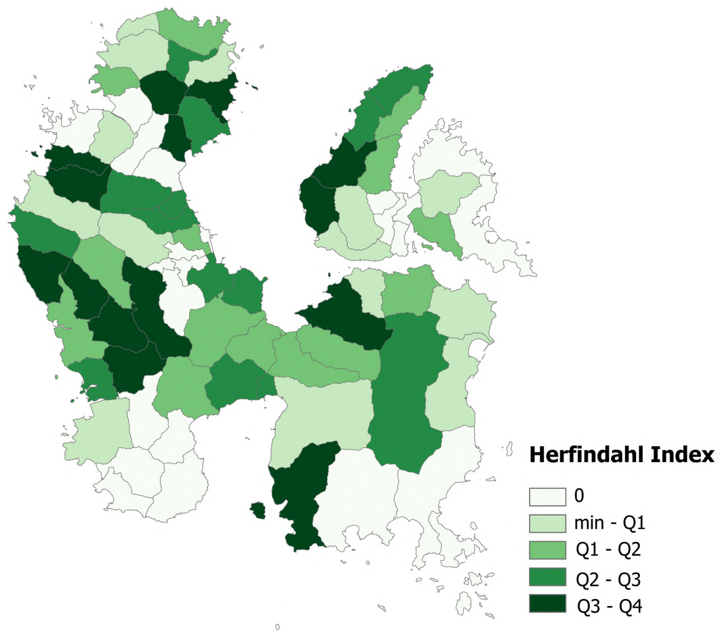
	Mean	SD	Min	Max
H_i	0.082	0.088	0	0.38
$TOPSHARE_i$	0.20	0.16	0	0.61
$CLANSHARE_i$	0.29	0.25	0	0.81

These proxies of village fractionalization have their shortcomings. A high clan density, per se, does not always imply lower fractionalization and a higher ability to self-organize. Some anecdotal evidence illustrates villages with high clan density that are highly fractionalized. A frequent example is when a few rival clans dominate a village. This would introduce noise to the proxy measures. However, I find evidence that the measures of clan density are positively correlated with more direct measures of village cohesiveness, such as the frequency of village meetings and public good provision. I discuss this further in Section 4: Mechanisms.

For the remainder of the paper, I define the villages above the median Herfindahl index as homogeneous villages and those below the median as heterogeneous villages.

first used it in Yang (2019)

Figure 2: Spatial Distribution of the Herfindahl Index in *Namhae* County



3 Homogeneity, Market Access, and Agricultural Productivity

My primary outcome of interest is agricultural productivity that I measure with two proxies: rice production per agricultural household and rice yield per hectare. Yield measures "land" productivity (Aragon et al. 2022), and output per agricultural household measures "farm" productivity.

Figure 3 motivates the analysis. It shows the distributions of rice output per agricultural household in 1969 and 1979. The dotted vertical bars indicate median values. On average, rice production grew by 70.7%, jumping from 0.83 tons/household to 1.38 tons/household. Figure A.10 shows that rice yield exhibited a similar increase, rising from 2.82 tons/hectare to 4.66 tons/hectare (a 65% increase). In addition to the rightward shift in the overall distribution, its spread also increased. The standard deviation increased from 0.38 tons in 1969 to 0.58 tons in 1979. Both the increase in the mean and the variance could reflect many factors. Some villages could have had more underutilized land or received more support from local governments.

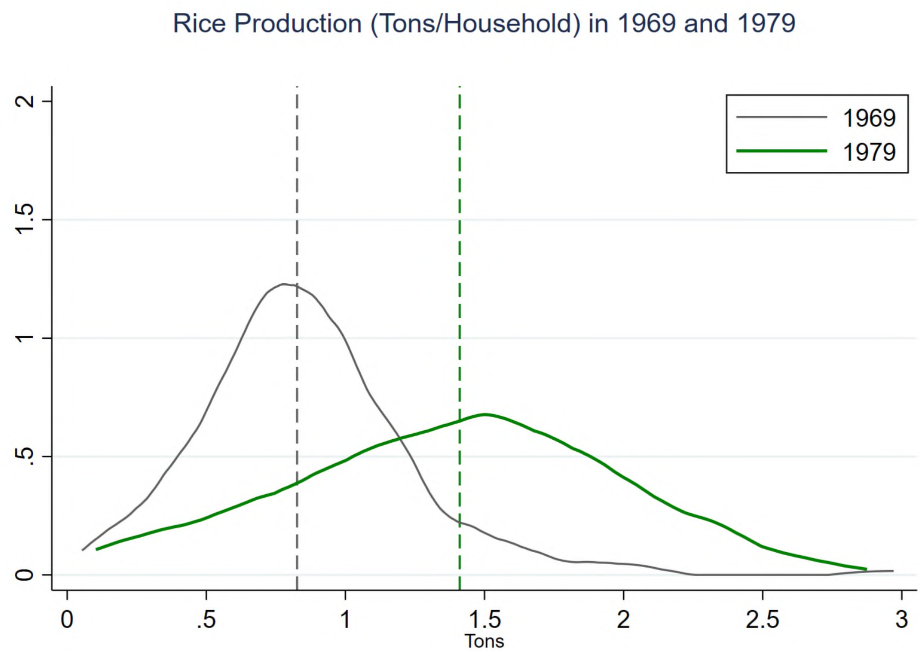


Figure 3: Evolution of Rice Production per Agricultural Household in Namhae County

To test if a village's clan homogeneity explains the increases in variation in agricultural productivity after the bridge construction, I employ a triple difference approach using heterogeneous villages and the nearby unconnected island as comparison groups. I then examine the dynamics of villages' responses to the bridge construction using an event study specification. Lastly, I look at heterogeneity by walking distance to the bridge within the connected island.

Triple Difference Estimate. The first difference compares the productivity between homogeneous villages (above the median Herfindahl Index) and heterogeneous villages (below the median). This difference is likely to be confounded by unobservable village and geographic characteristics. Economic shocks (such as the new bridge) that were taking place during this period could also confound this difference if clan homogeneity affects villages' ability to respond to them. I address this concern by using the nearby unconnected island as a control group. This island is particularly well-suited because it has always been a part of Namhae county, and therefore the quality of governance is comparable. As evident in Figure 1b, the two islands are located right next to each other, making market access before the bridge construction comparable. The population of the connected island was 104,359 (169 villages) in 1973, and the population of the unconnected island was 20,315 (26 villages).

The triple-difference (DDD) estimate of clan homogeneity is estimated by

$$\begin{aligned} Y_{it} = \alpha_i + \lambda_t + \beta_1 \cdot B_i \cdot H_i \cdot Post_t \\ + \beta_2 \cdot B_i \cdot Post_t + \beta_3 \cdot H_i \cdot Post_t + \epsilon_{it} \end{aligned} \tag{1}$$

In equation 1, Y_{it} denotes rice output per agricultural household in village i at year t; α_i is a village fixed effect; λ_t is a year fixed effect; B_i is an indicator for villages with access to the bridge; H_i is an indicator for the villages above the median Herfindahl index; and ϵ_{it} is the error terms, clustered at the village-level

The main parameter of interest is the triple difference estimate β_1. Interpreting β_1 as the causal estimate requires two assumptions. First is the standard parallel trends assumption that the differences between homogeneous and heterogeneous villages would have trended similarly in connected and unconnected islands in the absence of the bridge. I substantiate this assumption by checking that there are no differential pre-trends in the event study figures.

The second assumption is that there were no migrations between the two islands in anticipation of the bridge. This would be violated if, for example, people in the unconnected, homogeneous villages migrated to the connected island before the bridge was constructed. Though the data does not exist to check for selective migration between the islands, it is unlikely to be the case given the historical context. Homogeneous villages were historically formed because of immovable assets inherited from ancestors, such as a parcel of land or a house (The Academy of Korean Studies 1991). The inherited land often has tombs of ancestors, and for this reason, such assets are rarely traded. Village composition is not likely to have changed significantly in anticipation of the bridge.

Before discussing the estimation results, I present the summary statistics by treatment status and the relationship between the clan homogeneity measures and village/geographic characteristics.

Table 3 presents the means of the baseline village characteristics in 1972 by homogeneity and access to the bridge. I report the p-values for the comparison of the means. There

Table 3: Baseline Characteristics

	Clan Homogeneity			Access to the Bridge		
	Homogeneous	Heterogeneous	p-value	Connected	Unconnected	p-value
Baseline Characteristics (1972)						
Population	592.5	702.9	0.01	627.6	776.1	0.01
Households	101.9	131.9	0.02	109.7	163.0	0.19
Agri. Households	91.3	89.9	0.79	88.8	102.0	0.10
Share modified roofs	0.31	0.35	0.07	0.33	0.34	0.64
Share of villages with schools	0.16	0.08	0.08	0.22	0.19	0.72
Market Access (1972)						
Distance to the bridge (km)	22.7	25.5	0.12	24.1	N/A	N/A
Share of car accessible villages	0.78	0.73	0.52	0.77	0.65	0.26
Share of highway accessible villages	0.19	0.21	0.72	0.24	0	0.00
Share of state road accessible villages	0.20	0.18	0.72	0.16	0.42	0.02
Share of county road accessible villages	0.55	0.36	0.01	0.44	0.54	0.37
Agriculture (1969)						
Rice output per agri. household (kg/hh)	911.8	795.0	0.04	849.2	888.2	0.73
Rice yield (kg/ha)	1816.7	1588.0	0.01	1759.8	1344.0	0.00
Cultivated areas (ha)	46.2	45.0	0.65	44.0	56.3	0.01
Village Projects (1970-1979)						
Avg. budget per year (USD)	$3,972	$6,527	0.06	$5,150	$5,090	0.95
Share of budget on infrastructure	0.40	0.41	0.90	0.42	0.27	0.00
Distance to Administrative Units						
Distance to the closest admin office (km)	4.5	5.6	0.09	5.2	4.3	0.17
Distance to the county office (km)	9.4	13.1	0.00	11.3	N/A	N/A

Note: This table presents the means of the baseline characteristics of villages in the sample by clan homogeneity (in columns 1 and 2) and access to the bridge (in columns 4 and 5). A village is considered homogeneous if the Herfindahl index of clan concentration is greater than the median. Columns (3) and (6) display the p-values of the comparison of means across villages with and without homogeneity or bridge.

are some significant differences between the villages between the comparison groups. For example, homogeneous villages were smaller in population but had a higher rice yield and output per household. They were also closer to the public administration office and the county office. These are not surprising differences, as clans in homogeneous villages are likely to be descendants of former elite-class populations. Table A.11 reports the raw data of baseline village characteristics of the four comparison groups (homogeneous/heterogeneous villages; before/after the bridge).

I further test if the intensity of clan homogeneity is correlated with the village and

Table 4: Clan Homogeneity and Village Characteristics

	Population (1972)			Other Village Characteristics (1972)				
	(1) Share of agri. household	(2) Share of. pop. >14	(3) Share of women	(5) Age of leader	(6) Share of modified roofs	(7) log(num. of schools)	(8) 1[car accessibility]	(9) 1[access to water]
H_i	0.47**	0.02	0.01	1.19	-0.25*	-0.15	0.37	0.11
	(0.18)	(0.03)	(0.03)	(7.12)	(0.15)	(0.20)	(0.48)	(0.28)
$TOPSHARE_i$	0.32***	0.03	0.01	0.51	-0.11	-0.11	0.17	-0.08
	(0.10)	(0.02)	(0.01)	(3.69)	(0.08)	(0.12)	(0.26)	(0.15)
$CLANSHARE_i$	0.21***	0.02	0.00	-0.49	-0.04	-0.05	0.22	-0.03
	(0.06)	(0.01)	(0.01)	(2.63)	(0.06)	(0.08)	(0.15)	(0.10)
Observations	195	195	194	195	195	195	195	195
R-squared	~0.20	~0.00	~0.02	~0.00	~0.08	~0.10	~0.01	~0.01

Note: This table presents the relationship between three measures of clan homogeneity (H_i, $TOPSHARE_i$, and $CLANSHARE_i$) and the baseline village characteristics in 1972. Columns (1) to (3) show that clan homogeneity is associated with a greater share of households being agricultural, but no relationship with the share of population above 14 and the share of women. Columns (5) to (9) show that clan homogeneity is not associated with the following village characteristics: the age of a village leader, the share of modified roofs, the number of schools, car accessibility, or access to water. The village characteristics are from the *New Village Comprehensive Survey*, published by the government in 1972. The R-squared reports the average across three separate regressions using each measure of clan homogeneity. The coefficients are obtained from a regression of village characteristics on each measure of clan homogeneity controlling for the log of population. Standard errors are clustered at the 1930 village level.

geographic characteristics. Table 4 reports regression coefficients of village characteristics on three measures of clan homogeneity (H_i, $TOPSHARE_i$, and $CLANSHARE_i$), controlling for the log of population. Columns (1) to (3) show that clan homogeneity has a higher share of agricultural households but has no association with the share of the population above 14 and the share of women. Columns (5) to (9) show that clan homogeneity is also not associated with the village leader's age, the share of modified roofs, the number of schools, car accessibility, and access to water. These results are consistent across all three measures of clan homogeneity.

Taken together, homogeneous and heterogeneous villages differ in some baseline characteristics, but these differences do not vary with the intensity of clan homogeneity. However, it is important to note that the DDD estimation does not rely on the random allocation of the homogeneous villages. The identification comes from the bridge construction in 1973, and the DDD allows me to observe heterogeneity in villages' responses to the shock. In addition, the empirical strategy addresses the remaining identification concerns by controlling for time-invariant characteristics shared across the connected villages (B_i) and among the homogeneous villages (H_i). I also present the estimates with various sets of fixed effects: village, time, township-year, and farmland-year fixed effects. I also account for the baseline differences with an extensive set of controls for geographic and village characteristics.

The triple-difference estimates based on equation 1 are presented in Table 5. The estimates with (1) no controls, (2) with demographic controls, and (3) with demographic and village-level controls suggest the impact of clan homogeneity and access to the bridge to be 280 kgs of additional rice output per household. Including geographic controls in column (4) increases the estimate to 330 kgs. The household's average rice output during this sample period was 1,135 kg, which implies that the connected, homogeneous villages experienced 25% to 30% faster productivity growth.

Table 6 shows that the estimates are robust to alternative sets of fixed effects. The estimates are consistent when including (1) only the village-fixed effects, (2) the village and year fixed effects, (3) the village, year, and township-year fixed effects, and (4) the village, year, and the 1969 farmland area-year fixed effects. Township is an administrative unit above the village, with about 25 villages under its jurisdiction. The township-year fixed effects control for the fact that agricultural productivity was trending upwards but at different rates by townships. The 1969 farmland area-year fixed effects provide additional flexibility by allowing productivity to evolve differently depending on villages' initial endowments of farmland.

Lastly, I present the DDD estimates using alternative measures of clan homogeneity (TOPSHARE and CLANSHARE) in Table A.12. The estimates are consistent across all three measures.

Table 5: The DDD Estimates of Clan Homogeneity on Agricultural Productivity

	Dependent variable: Rice Output/Agri. Household			
	(1)	(2)	(3)	(4)
Constant	1,031.32***	2,519.52***	2,159.78***	2,747.43***
	(94.49)	(643.72)	(599.19)	(551.39)
Homogeneity	-157.95	-257.37**	-347.40**	-343.29***
	(119.18)	(122.63)	(138.77)	(119.18)
Bridge	-299.85***	-412.58***	-448.24***	-370.25***
	(103.02)	(113.64)	(125.30)	(112.38)
Homogeneity × Bridge	340.92***	323.45**	389.48***	371.02***
	(130.80)	(132.58)	(149.43)	(131.91)
post	392.96***	419.48***	410.03***	456.00***
	(106.17)	(117.85)	(118.85)	(136.67)
Homogeneity × post	-229.93*	-228.79*	-225.08*	-273.74**
	(122.09)	(122.45)	(123.45)	(137.82)
Bridge × post	-42.00	-52.65	-46.83	-102.87
	(113.41)	(126.21)	(123.85)	(138.93)
Homogeneity × Bridge × post	**285.01****	**279.97****	**286.79****	**330.55****
	(130.90)	**(134.32)**	**(135.06)**	**(144.89)**
Mean of Dep. Var.	1135	1135	1135	1135
Demographic controls	No	Yes	Yes	Yes
Socioeconomic/village controls	No	No	Yes	Yes
Geographic controls	No	No	No	Yes
Observations	1,338	1,332	1,332	1,217
R-squared	0.13	0.21	0.25	0.37

Note: The demographic controls include the log of population, the share of agricultural households, the share of the population greater than 14, and the share of women. The socioeconomic/village controls include the age of a village leader, the share of modified roofs, log of the number of schools, car accessibility, access to water, electricity, and phone. The geographic controls include log of total land area, share of land area that is cultivated, and log of distance to the closest public office. These are the full set of variables for which descriptive statistics are provided in Table A.11. Standard errors are clustered at the village level.

Table 6: The DDD Estimates of Clan Homogeneity on Agricultural Productivity

	Dependent variable: Rice Output/Agri. Household			
	(1)	(2)	(3)	(4)
Homogeneity × Bridge × post	**317.10****	**362.73****	**344.70****	**359.40****
	(143.90)	**(148.08)**	**(146.25)**	**(159.72)**
Mean of Dep. Var.	1135	1135	1135	1135
Village FEs	Yes	Yes	Yes	Yes
Year FEs	No	Yes	Yes	Yes
Township-Year FEs	No	No	Yes	No
Farmland Area-Year FEs	No	No	No	Yes
Observations	1,217	1,217	1,217	1,202
R-squared	0.68	0.78	0.83	0.80

Note: This table shows the estimates with alternative sets of fixed effects: including (1) only the village-fixed effects, (2) the village and year fixed effects, (3) the village, year, and township-year fixed effects, and (4) the village, year, and the 1969 farmland area-year fixed effects. Township is an administrative unit above the village, with about 25 villages under its jurisdiction. The township-year fixed effects control for the fact that agricultural productivity was trending upwards but at different rates by townships. The 1969 farmland area-year fixed effects provide additional flexibility by allowing productivity to evolve differently depending on villages' initial endowments of farmland.

Event Studies. I turn to event studies to illustrate the dynamics of villages' responses. The event study specification is:

$$Y_{it} = \alpha_i + \lambda_t + \sum_{\tau \neq 1972} \beta_\tau \cdot B_i \cdot H_i \cdot 1[t=\tau] + \sum_{\tau \neq 1972} \beta_\tau \cdot B_i \cdot 1[t=\tau] + \sum_{\tau \neq 1972} \beta_\tau \cdot H_i \cdot 1[t=\tau] + \epsilon_{it} \tag{2}$$

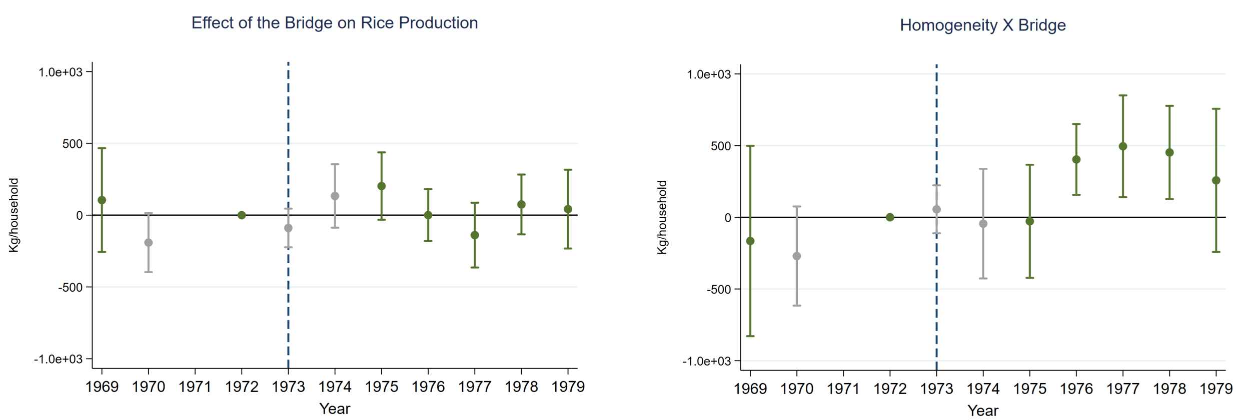
The notation is the same as in equation 1. I first look at the effect of the bridge construction alone, comparing the connected and unconnected islands before and after the new bridge. Then, I allow heterogeneity in responses to the new bridge by estimating equation 2.

Figure 4a plots the event study of rice productivity comparing connected and unconnected villages. I use the grey color to mark the years with an incomplete sample (1970, 1973, and 1974). For these three years, data only exists for 49 villages out of 195 total villages. The year 1971 is empty because data does not exist; I omit the year 1972 using it as the baseline. Rice production seems to increase with the new bridge, but the effect is not statistically significant.

Figure 4b plots the event study coefficients from equation 2, observing heterogeneity by clan density. First, connected, homogeneous villages do not seem to be on differential pre-trends. This supports the parallel trends assumption. Second, This figure also reveals an interesting dynamic in villages' responses. Though the bridge was constructed in 1973, the productivity gap did not manifest until 1976. This gap is also short-lived. Figure A.11 shows the results from the same specification for a longer time frame. I further investigate the driving forces behind the lag in Section 4: Mechanisms. I find that faster technology adoption in connected, homogeneous villages (primarily mechanized tillers) starting in 1976 could explain this productivity gap. I provide suggestive evidence that in the years between the bridge construction in 1973 and 1976, these villages were making investments in public goods, such as village and farm roads, that were necessary to operate the mechanized tillers.

I use an alternative measure of rice productivity—rice yield—and get similar results. The same event study specification using rice yield is presented in Figure A.12.

Figure 4: Rice Productivity and the Construction of the *Namhae* Bridge (1973)



(a) The Effect of the Bridge

(b) Homogeneity and the Bridge

Note: Difference in rice productivity for (a) connected and unconnected villages, and (b) heterogeneity by clan density around the bridge construction is plotted. The coefficients are estimated using equation 2. Standard errors are clustered at the village level.

Using Distance to the Bridge. I turn to heterogeneity of the treatment effect by walking distance to the bridge among the villages in the connected island. Given that feeder and rural roads were not paved in the 70s, remote villages were exposed to uncertain road conditions. Some villages were as far as 52 km away from the bridge. Therefore, walking distance proxies the intensity of the bridge shock.

I estimate the travel distances between the villages and the bridge in 1970 using the distance matrices published in the 1970 Statistical Yearbook of Namhae County. Using contemporary travel distances or Euclidean distances can be misleading. Not only had more direct routes been built since the 1970s, but the connected island also has mountainous terrains with the villages naturally forming near seashores. I discuss the construction of the distance variable in detail in Appendix A.2, and the summary statistics are provided in Table A.13.

Figure 5 motivates the analysis by providing visual evidence for variation in productivity growth. It plots differences in rice productivity between 1969 and 1979. The villages closer to the bridge seem to have experienced faster productivity growth than those farther away.

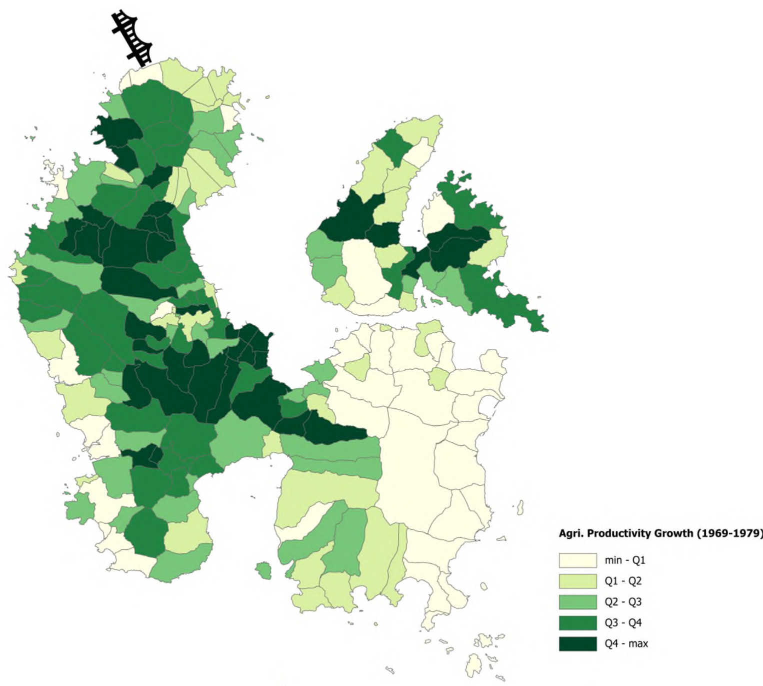


Figure 5: Agricultural Productivity Growth between 1969 and 1979

Table 7 presents the DDD estimation results based on equation 1, where I replace B_i with $LongDistance_i$, which equals 1 when a village is farther away than the median distance to the bridge (23.5 km). The results suggest that the DDD estimates presented in Table 6 are driven by homogeneous villages in the connected island that are far from the bridge. As the distance to the bridge increased (and the intensity of the bridge decreased), homogeneous villages experienced higher agricultural productivity growth compared to heterogeneous villages.

Table 7: The DDD Estimates of Clan Homogeneity on Agricultural Productivity

	Dependent variable: Rice Output/Agri. Household			
	(1)	(2)	(3)	(4)
Homogeneity × Long Distance × post	**306.47*** **	**262.56*** **	**292.42*** **	**300.34*** **
	(89.63)	**(90.24)**	**(89.83)**	**(80.56)**
Mean of Dep. Var.	1135	1135	1135	1135
Village FEs	Yes	Yes	Yes	Yes
Year FEs	No	Yes	Yes	Yes
Township-Year FEs	No	No	Yes	No
Farmland Area-Year FEs	No	No	No	Yes
Observations	953	953	953	943
R-squared	0.74	0.84	0.90	0.86

Note: *LongDistance* is an indicator that equals 1 when a village is farther away than the median distance to the bridge (23.5 km). This table shows the estimates with alternative sets of fixed effects: including (1) only the village-fixed effects, (2) the village and year fixed effects, (3) the village, year, and township-year fixed effects, and (4) the village, year, and the 1969 farmland area-year fixed effects. Township is an administrative unit above the village, with about 25 villages under its jurisdiction. The township-year fixed effects control for the fact that agricultural productivity was trending upwards but at different rates by townships. The 1969 farmland area-year fixed effects provide additional flexibility by allowing productivity to evolve differently depending on villages' initial endowments of farmland.

Table 8 corroborates the results by showing (1) a double difference looking only at the effect of being far away, (2) a double difference looking only at the effect of clan homogeneity, and (3) the DDD estimates. While the overall productivity of the remote villages improved less than the close-by villages, homogeneous places experienced faster productivity growth than heterogeneous villages.

Table 8: The Estimates of Clan Homogeneity on Agricultural Productivity

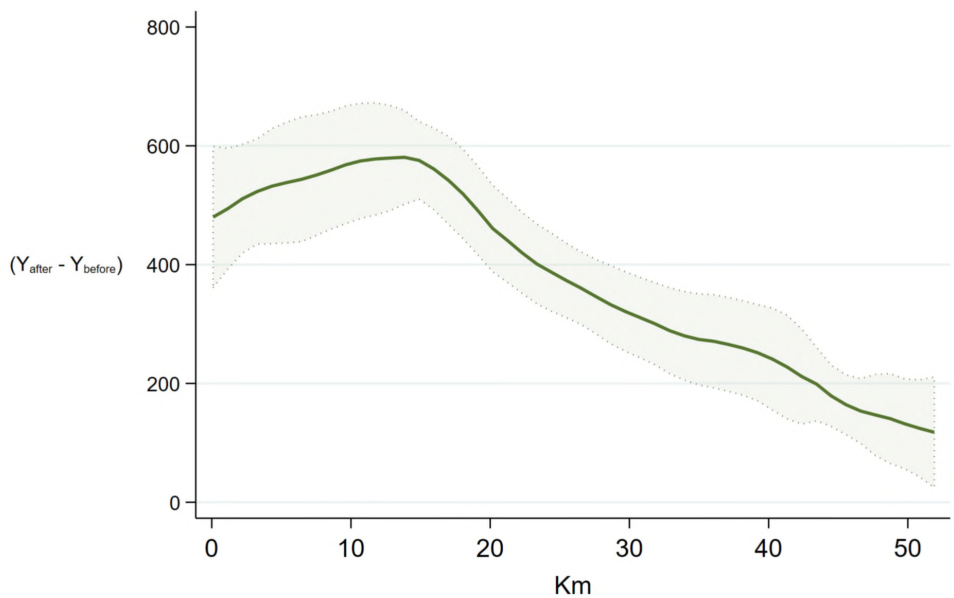
	Rice Output/Agri. Household		
	(1)	(2)	(3)
Long Distance × post	-219.19***		-386.70***
	(38.10)		(61.32)
Homogeneity × post		42.75	-163.63**
		(44.26)	(66.53)
Homogeneity × Long Distance × post			300.34***
			(80.56)
Mean of Dep. Var.	1135	1135	1135
Village FEs	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Farmland Area-Year FEs	Yes	Yes	Yes
Observations	943	943	943
R-squared	0.86	0.85	0.86

Note: This table shows the estimates with (1) a double difference comparing close-by and remote villages, (2) homogeneous and heterogeneous villages, and (3) the DDD differences that compare four groups (homogeneous/heterogeneous, close-by/remote villages). All regressions include village-fixed effects, year-fixed effects, and the 1969 farmland area-year fixed effects.

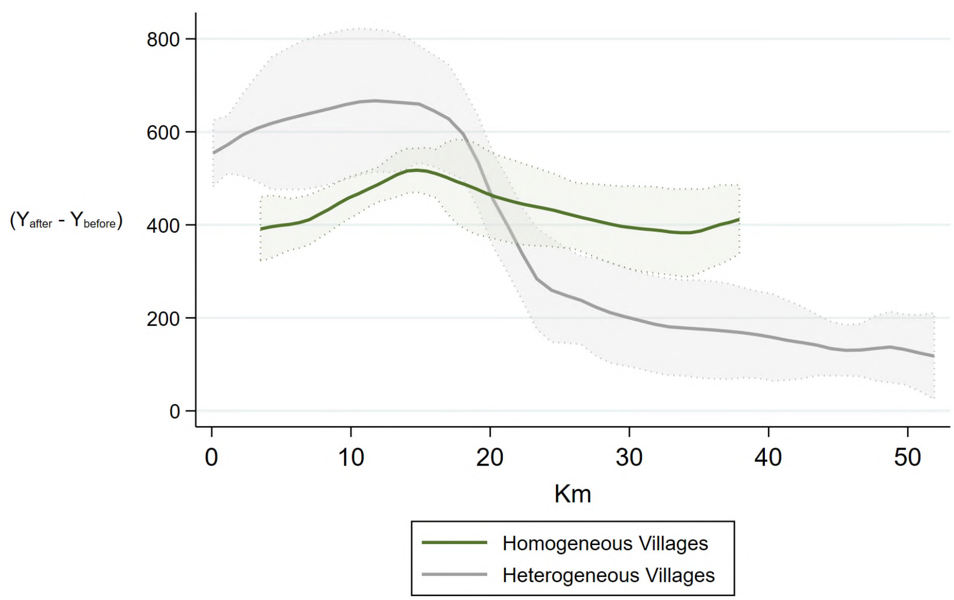
I explore these results further by plotting the difference in average rice output per agricultural household before and after the bridge ($\bar{Y}_{before} - \bar{Y}_{after}$) nonparametrically as a function of distance to the bridge. The top panel of Figure 6a presents the plot for all villages in the connected island. Overall, the gains exhibit a hump-shaped pattern, in which the gains were highest for the villages 15 km away from the bridge and steadily decreased with distance. This pattern could reflect that the closest villages might have already had good access to maritime transportation so the marginal effect of the bridge on productivity is lower. On the other hand, the effect could be more significant for those that were 15 km away that previously rely on both ground and maritime transportation.

The bottom panel of Figure 6b decomposes the top panel's figure by clan homogeneity. Heterogeneous villages experienced a steeper decline in productivity gains with distance, whereas the pattern is flatter for homogeneous villages. This is consistent with the results from Table 8.

Figure 6: Non-parametric Estimates of Rice Productivity Growth by Distance to the Bridge



(a) Nonparametric Estimation of Productivity Growth



(b) Nonparametric Estimation of Productivity Growth by Homogeneity

Note: The figures plot local polynomial regression of the average productivity growth ($\bar{Y}_{before} - \bar{Y}_{after}$) on distance to the bridge. The top panel includes all connected villages; the bottom panel decomposes it by homogeneity.

Taken together, after the bridge construction in 1973, homogeneous villages experienced about 30% faster agricultural productivity growth between 1969 and 1979 than heterogeneous villages. The event study illustrates no differential pre-trends and a three-year lag until the effect manifested. Lastly, looking at treatment heterogeneity by distance to the bridge reveals that the effect was driven by homogeneous villages that were remote from the bridge. While the heterogeneous villages experienced a faster decline in productivity gains with increasing travel distance, homogeneous villages gained uniformly regardless of the distance.

4 Mechanisms

I turn to investigating the mechanisms that helped homogeneous villages experience faster productivity growth. I also provide suggestive evidence that could explain the lag between the bridge construction and when the effect on productivity appears.

Technology Adoption. Throughout the 1970s, the Korean government campaigned for adopting "scientific approaches to agriculture practices" (The Ministry of Home Affairs 1980). A particular emphasis was put on expanding the adoption of mechanical tillers. A tiller's main function is to stir and pulverize the soil before planting, but it can also be customized to perform other tasks such as transporting materials or threshing grains. Tillers are still widely used as multi-functional vehicles in rural regions of Korea. According to the agricultural census, the share of households with mechanical tillers climbed from 0.3% in 1970 to 10.5% in 1980.

I evaluate if tiller adoption varied by bridge-connectedness and clan homogeneity, using equation 1. Namhae county started widely tracking tiller ownership in 1976, which was after the bridge construction in 1973. This makes the comparison before and after the bridge impossible. To circumvent this issue, I assume that no household owned a tiller in 1972, a year prior to the bridge construction. This assumption is reasonable given that there was only an average of 2.8 tillers in a village in 1976. The number steadily increased to 3.7 tillers in 1977, 4.5 tillers in 1978, and 5.6 tillers in 1979.

Figure 7 shows the dynamics of tiller ownership until 1984. The households in homogeneous, connected villages adopted new technology faster. Importantly, the difference becomes statistically significant in 1976, which coincides with the year that the productivity gap ap-

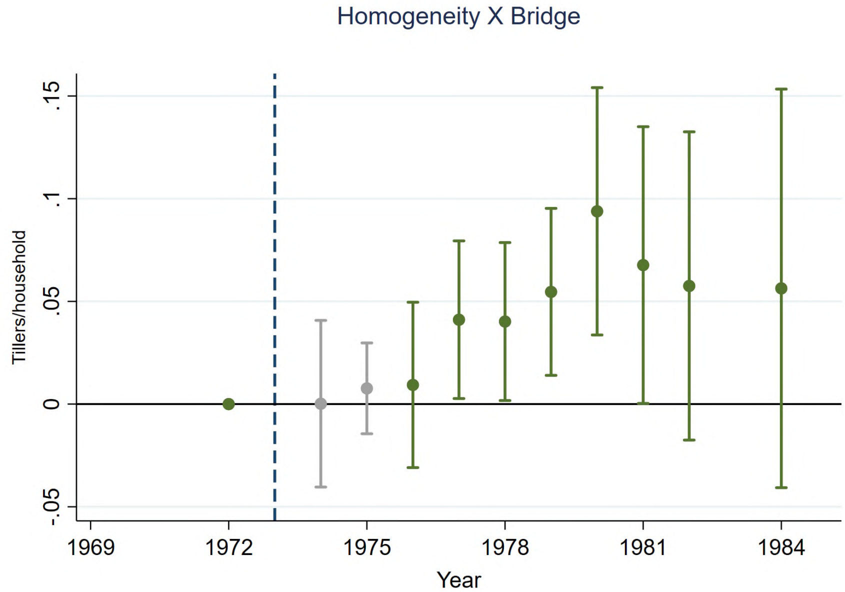


Figure 7: The DDD Estimation of Tiller Ownership between 1969 and 1979

peared in Figure 4b.

I further investigate how the tiller adoption varied by the intensity of clan homogeneity. Figure 8 shows a scatter plot of the Herfindahl index of clan concentration in the connected island on the x-axis and the number of tillers per household averaged between 1976 and 1979 on the y-axis. The pattern is consistent with the underlying hypothesis that homogeneous villages adopted new technology faster, which could have driven their productivity growth.

I corroborate my findings with the home-production and the use of fertilizers. Figure 9 shows the event study coefficients of home-produced fertilizers, based on equation 1. The data for fertilizers is not available for the years 1970, 1971, and 1973, and is incomplete for the years 1974 and 1975 (in grey colors). Though I do not observe a sharp increase in 1976 and the results are not statistical for all years, the figure shows that connected, homogeneous villages produced more fertilizers during the sample period.

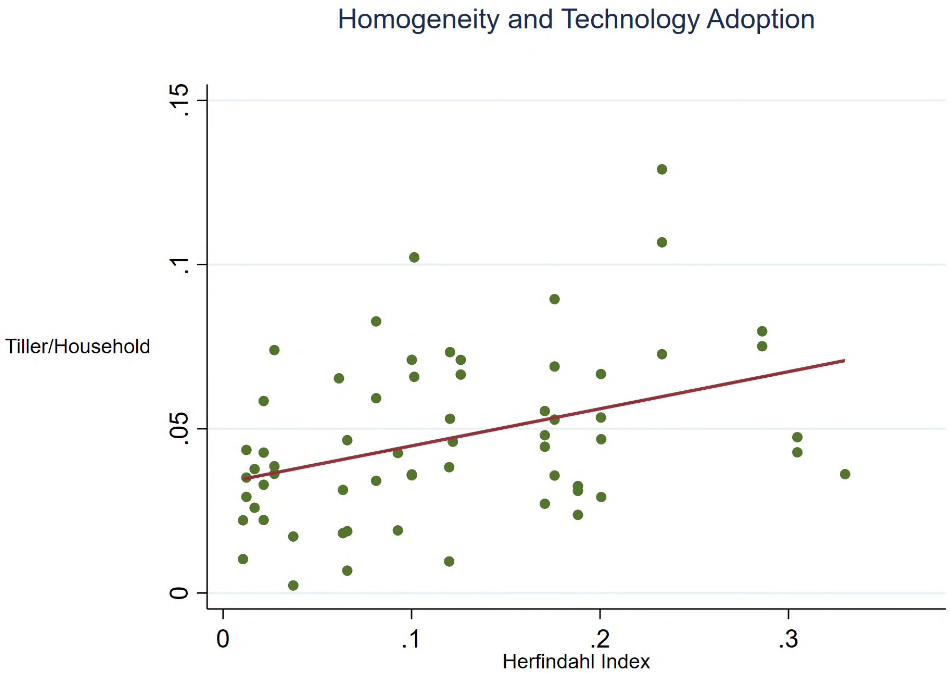


Figure 8: Tiller Ownership by Clan Homogeneity

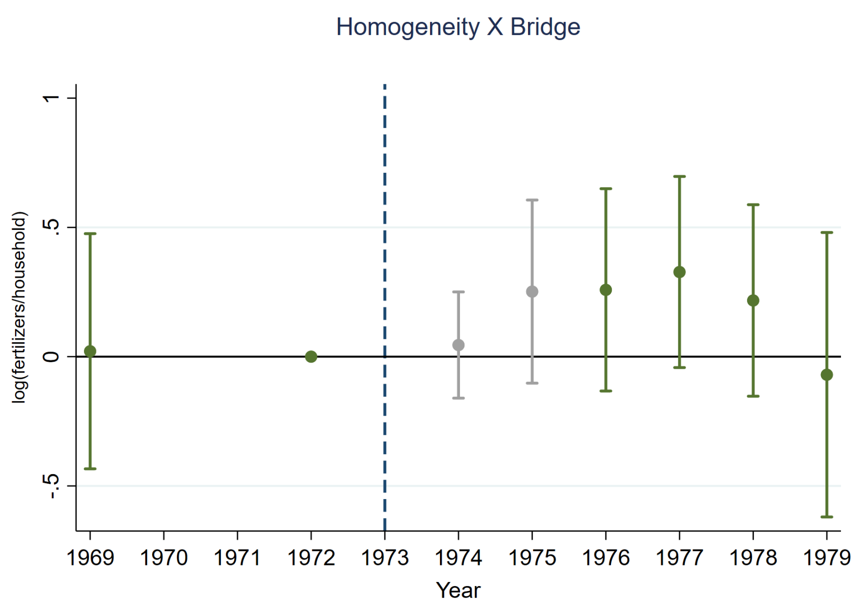


Figure 9: The DDD Estimation of Home-produced Fertilizers between 1969 and 1979

Public Good Provision. Why were connected, homogeneous villages able to adopt the new technology faster? I provide evidence that homogeneous villages were better able to self-organize for providing public goods, such as village and farm roads. Given the road conditions at the time, public good investments were necessary to use tillers. The demand for better roads is likely to have increased, given the new economic opportunities (better market access thanks to the bridge) and availability of new technology (tillers). Yet, homogeneous villages took the lead in exploiting them because they were able to promptly make complementary investments. I provide suggestive evidence that supports this claim. I show that connected, homogeneous villages were promoted to the highest grade of A, which measured their abilities to provide public goods. I substantiate this claim by showing a positive correlation between the intensity of clan homogeneity and direct measures of social cohesion, such as the frequency of village meetings and the number of meeting participants per household.

First, Figure 10 plots the event study results for an indicator of whether a village received a grade of A from the annual evaluation. Every village was evaluated annually for their performance on the public good provision and received one of the three-tiered grades (A, B, or C). The first evaluation was published in 1972, and thus, I do not observe a pre-trend. However, the connected, homogeneous villages seem to have outperformed in the years following the bridge construction. The negative coefficients between 1976 and 1979 are mechanical, as grade A was the highest grade a village could receive. This pattern shows that other villages caught up in the subsequent years.

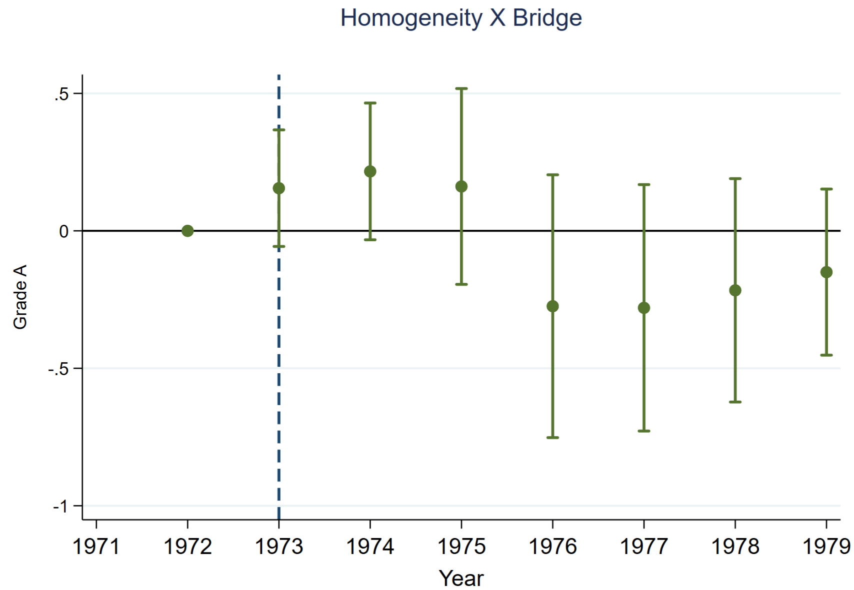


Figure 10: The DDD Estimation of Village Grades between 1972 and 1979

I provide further evidence that homogeneous villages made more agricultural investments that were non-excludable and non-rivalrous in nature. Table 10 shows the relationship between three alternative measures of clan homogeneity and public good provision. Clan homogeneity is associated with (1) getting promoted to grade A earlier, (2) building more village roads, (3) a higher voluntary participation rate in village projects, (4) a higher reservoir capacity, and (5) home-made fertilizers. These patterns are all consistent with the hypothesis that public good provision was a channel through which clan homogeneity affected agricultural productivity.

Table 9: Clan Homogeneity and Public Good Provision

	Public Good Provision				
	(1) Years with grade A	(2) log(total length of village roads/hh)	(3) log(# of project participants/hh)	(4) Reservoir capacity per irrigated land	(5) log(home-produced fertilizers/hh)
H_i	3.72**	4.13***	4.84**	1.47**	1.33***
	(1.76)	(1.44)	(1.90)	(0.60)	(0.44)
$TOPSHARE_i$	1.75*	2.01**	2.29**	0.82**	0.84***
	(0.97)	(0.80)	(1.08)	(0.33)	(0.28)
$CLANSHARE_i$	1.20*	1.56***	1.83***	0.44*	0.43**
	(0.66)	(0.51)	(0.69)	(0.24)	(0.19)
Mean of dept. var.	3.30 years	2.03	2.89	0.44 tons/ha	2.32
Observations	196	196	196	196	196
R-squared	∼0.02	∼0.03	∼0.03	∼0.04	∼0.09

Note: This table presents the relationship between three measures of clan homogeneity (H_i, $TOPSHARE_i$, and $CLANSHARE_i$) and public good provision. Columns (1) to (5) show that homogeneous villages provision more public goods. I use measurements of public good provision for years between 1969 and 1979. The R-squared reports the average across three separate regressions using each measure of clan homogeneity. The coefficients are obtained from an univariate regression of each measure of clan homogeneity with the measures of social cohesion, with standard errors clustered at the 1930 village-level.

Lastly, this is also reflected in village meetings. Table 10 shows that homogeneous villages (1) have more meeting logs recorded in their village history books, (2) have more people show up to the meetings, but (3) the meetings have a smaller share of female participants. This is consistent with the argument that homogeneous villages are more traditional and therefore cohesive, but also patriarchal.

Table 10: Clan Homogeneity and Social Cohesion

	Village Meetings		
	(1)	(2)	(3)
	# of recorded meetings/hh	# of participants per household	Share of female participants
H_i	0.16**	0.60***	-0.24**
	(0.07)	(0.22)	(0.10)
$TOPSHARE_i$	0.06	0.33**	-0.13**
	(0.04)	(0.14)	(0.06)
$CLANSHARE_i$	0.05**	0.16*	-0.10**
	(0.02)	(0.08)	(0.04)
Mean of dept. var.	0.05	0.70 people/hh	0.19
Observations	196	103	108
R-squared	~0.01	~0.05	~0.05

Note: This table presents the relationship between three measures of clan homogeneity (H_i, $TOPSHARE_i$, and $CLANSHARE_i$) and the direct measures of social cohesion. Columns (1) to (3) show that more meetings are recorded, and the participation rate is higher in homogeneous villages, but the share of women in the meetings is lower. I use measurements of village meetings from all available data in the *History of Our Village*. The R-squared reports the average across three separate regressions using each measure of clan homogeneity. The coefficients are obtained from an univariate regression of each measure of clan homogeneity with the measures of village meetings, with standard errors clustered at the 1930 village-level. The sample sizes for columns (2) and (3) are smaller because some village did not record the number of participants.

5 Conclusion

Korea's rural development in the 1970s provides a unique opportunity to study how social fractionalization affects a group's ability to respond to economic shocks. Its unique social context isolates one dimension of social fractionalization (family clan homogeneity) that is otherwise complex and difficult to measure. Additionally, the Korean government pursued ambitious national infrastructure projects that created time variations in economic opportunities. The government also adopted a decentralized approach to rural development, delegating the authority for public good provision to local communities. This policy made local, informal governance more relevant for rural development. And most importantly, Korea experienced a 60% growth in agricultural output during this period.

Using the construction of the *Namhae* bridge as an economic shock, I find that less fractionalized villages, proxied by clan homogeneity, saw better economic outcomes. Homogeneous villages experienced about 30% faster agricultural productivity growth between 1969 and 1979. These differences were driven by faster technology adoption and better public good provision. Technology adoption improved agricultural productivity, and the village-level investments in complementary public goods made such a fast pace of technology adoption possible.

Taken together, the results suggest that social fractionalization influences local-level responses to new economic opportunities and policies. My results underpin a long literature on the importance of historical institutions in economic development.

These results also help explain the success of Korea's rural development. Homogeneous villages promptly responded to new economic opportunities, and the decentralized, bottom-up approach to rural development supported such responses.

There are multiple ways that this paper can be extended. While its analysis is limited to a county consisting of two islands, the Korean government rapidly expanded the road network throughout the 1970s. This naturally created variations in economic opportunities nationwide. Additionally, the mechanism of the catch-up by heterogeneous villages has not been identified. It is not clear whether the catch-up was driven by the demonstration effect (i.e., proving the feasibility of village projects and their potential economic impact) or by spillovers (i.e., investing in a village's road network creating positive externalities to nearby villages). A natural follow-up question is whether this kind of fast agricultural productivity growth would have been attainable in the absence of homogeneous groups. I leave these agenda for future research.

References

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A Olken, and Julia Tobias, "Targeting the poor: evidence from a field experiment in Indonesia," *American Economic Review*, 2012, *102* (4), 1206–40.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg, "Fractionalization," *Journal of Economic growth*, 2003, *8* (2), 155–194.

Aragon, Fernando M, Diego Restuccia, and Juan Pablo Rud, "Are small farms really more productive than large farms?," *Food Policy*, 2022, *106*, 102168.

Asher, Sam and Paul Novosad, "Rural roads and local economic development," *American Economic Review*, 2020, *110* (3), 797–823.

Bossert, Walter, Conchita d'Ambrosio, and Eliana La Ferrara, "A generalized index of fractionalization," *Economica*, 2011, *78* (312), 723–750.

Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padró i Miquel, "The value of democracy: evidence from road building in Kenya," *American Economic Review*, 2015, *105* (6), 1817–51.

Esteban, Joan and Debraj Ray, "Linking conflict to inequality and polarization," *American Economic Review*, 2011, *101* (4), 1345–74.

_ , Laura Mayoral, and Debraj Ray, "Ethnicity and conflict: An empirical study," *American Economic Review*, 2012, *102* (4), 1310–42.

Gebresilasse, Mesay, "Rural roads, agricultural extension, and productivity," 2021.

Han, Do-Hyun, "1970년대 새마을운동에서 마을 지도자들의 경험세계: 남성 지도자들을 중심으로 [Life World of Village Leaders during the Saemaul Movement in the 1970s: Focusing on Male Leaders]," *사회와역사 [Society and History]*, 2010, *88*, 267–305.

Heß, Simon, Dany Jaimovich, and Matthias Schündeln, "Development projects and economic networks: Lessons from rural gambia," *The Review of Economic Studies*, 2021, *88* (3), 1347–1384.

Hjort, Jonas, "Ethnic divisions and production in firms," *The Quarterly Journal of Economics*, 2014, *129* (4), 1899–1946.

Korea National Statistics Office, "Gross Value Added of Agriculture, Forestry, and Fisheries," https://kosis.kr/eng/ 2022.

Moscona, Jacob, Nathan Nunn, and James A Robinson, "Segmentary Lineage Organization and Conflict in Sub-Saharan Africa," *Econometrica*, 2020, *88* (5), 1999–2036.

Namhae County Office, *Statistical Yearbook of Namhae County*, Namhae County, 1974.

Shamdasani, Yogita, "Rural road infrastructure & agricultural production: Evidence from India," *Journal of Development Economics*, 2021, *152*, 102686.

Shen, Zejiang, Kaixuan Zhang, and Melissa Dell, "A large dataset of historical japanese documents with complex layouts," in "Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops" 2020, pp. 548–549.

The Academy of Korean Studies, *Encyclopedia of Korean Culture*, Vol. 28, DongBang Media Co., 1991.

The Ministry of Home Affairs, *새마을운동10년사 [Ten Year History of New Village Movement]*, Seoul: The Ministry of Home Affairs, 1980.

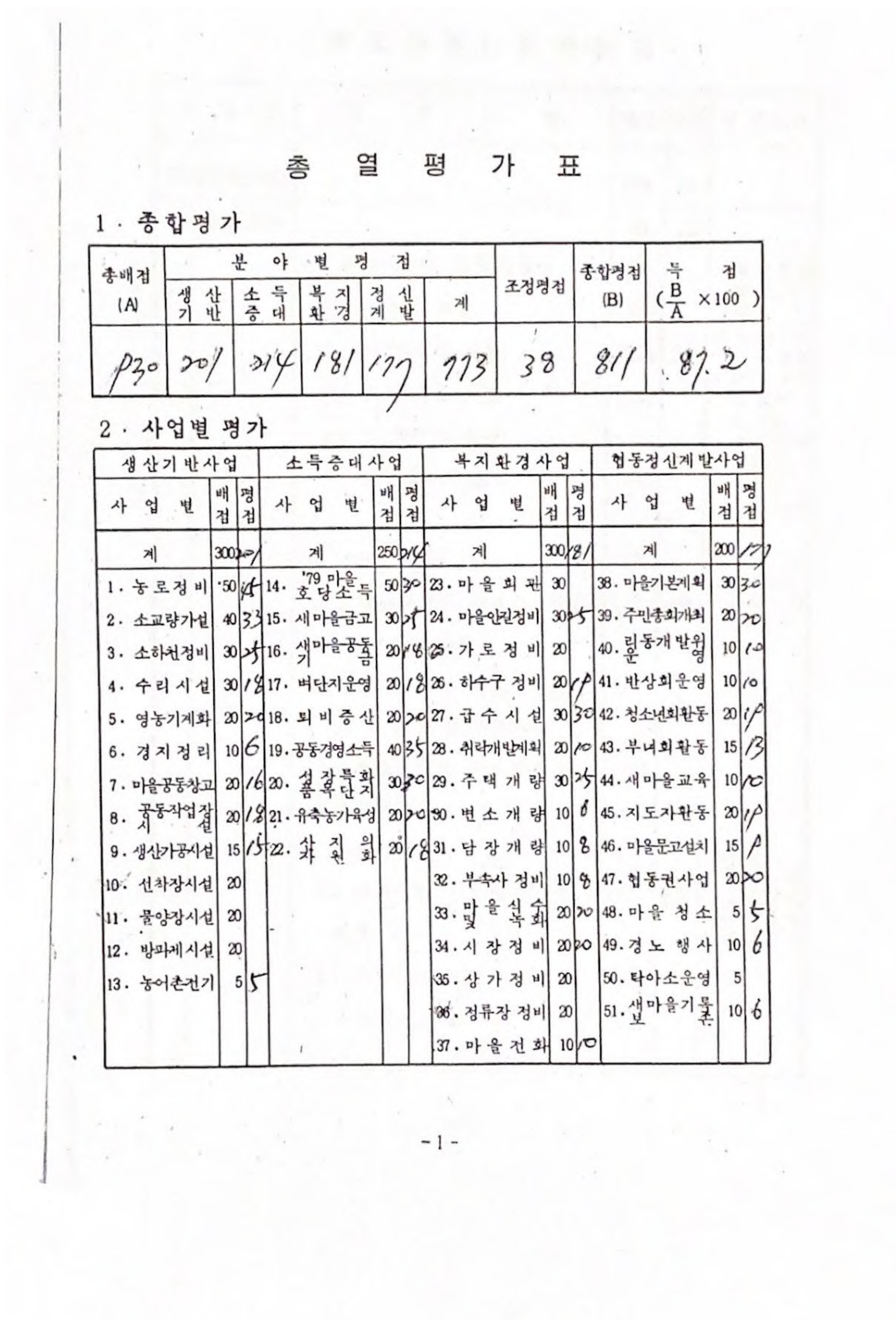
Yang, Hyunjoo, "Family clans and public goods: Evidence from the New Village Beautification Project in South Korea," *Journal of Development Economics*, 2019, *136*, 34–50.

A Online Appendix

Contents

A.1 Additional Figures and Tables

Figure A.1: An Example Page of Annual Village Evaluations



종 열 평 가 표

1 · 종합평가

총배점 (A)	분야별평점: 생산기반	소득증대	복지환경	정신계발	계	조정평점	종합평점 (B)	득점 ($\frac{B}{A}\times 100$)
930	201	214	181	177	773	38	811	87.2

2 · 사업별 평가

생산기반사업: 사업별	배점	평점	소득증대사업: 사업별	배점	평점	복지환경사업: 사업별	배점	평점	협동정신계발사업: 사업별	배점	평점
계	300	201	계	250	214	계	300	181	계	200	177
1. 농로정비	50	45	14. '79 마을 호당소득	50	30	23. 마을회관	30		38. 마을기본계획	30	30
2. 소교량가설	40	33	15. 새마을금고	30	25	24. 마을안길정비	30	25	39. 주민총회개최	20	20
3. 소하천정비	30	25	16. 새마을공동기금	20	18	25. 가로정비	20		40. 리동개발위 운영	10	10
4. 수리시설	30	18	17. 벼단지운영	20	18	26. 하수구 정비	20	18	41. 반상회운영	10	10
5. 영농기계화	20	20	18. 퇴비증산	20	20	27. 급수시설	30	30	42. 청소년회활동	20	18
6. 경지정리	10	6	19. 공동경영소득	40	35	28. 취락개발계획	20	10	43. 부녀회활동	15	13
7. 마을공동창고	20	16	20. 설장특화 품목단지	30	30	29. 주택개량	30	25	44. 새마을교육	10	10
8. 공동작업장 시설	20	18	21. 유축농가육성	20	20	30. 변소개량	10	8	45. 지도자활동	20	18
9. 생산가공시설	15	15	22. 삼지의 자원화	20	18	31. 담장개량	10	8	46. 마을문고설치	15	8
10. 선착장시설	20					32. 부속사 정비	10	8	47. 협동권사업	20	20
11. 물양장시설	20					33. 마을식수 및 녹화	20	20	48. 마을청소	5	5
12. 방파제시설	20					34. 시장정비	20	20	49. 경노행사	10	6
13. 농어촌전기	5	5				35. 상가정비	20		50. 탁아소운영	5	
						36. 정류장 정비	20		51. 새마을기록 보존	10	6
						37. 마을전화	10	10			

- 1 -

Figure A.2: Village Roads and Tillers in the 1970s

Figure A.3: Agricultural Productivity Growth in Korea in the 1970s

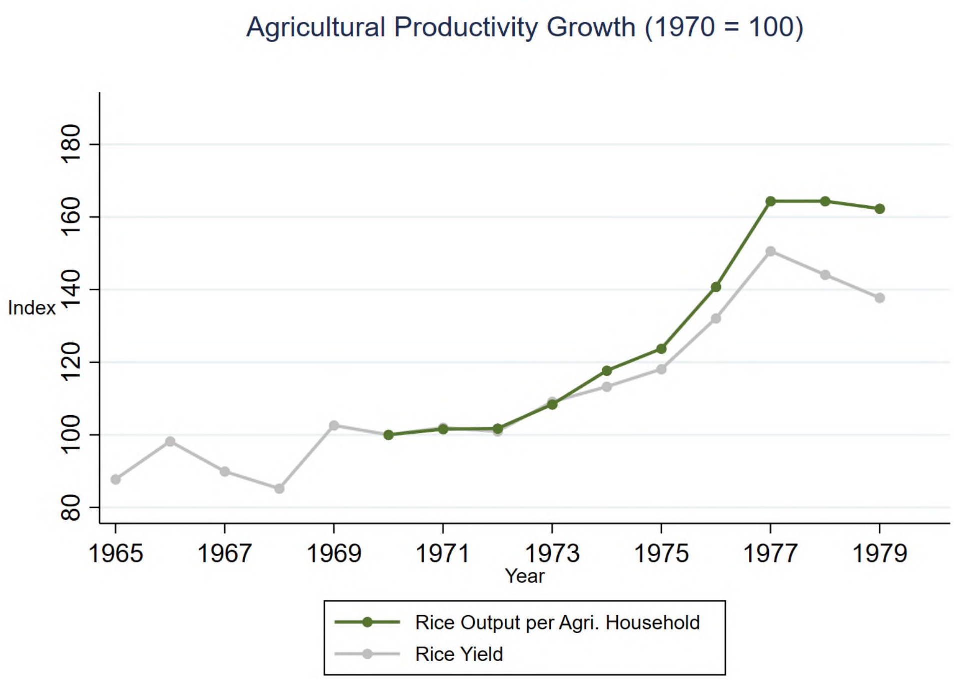
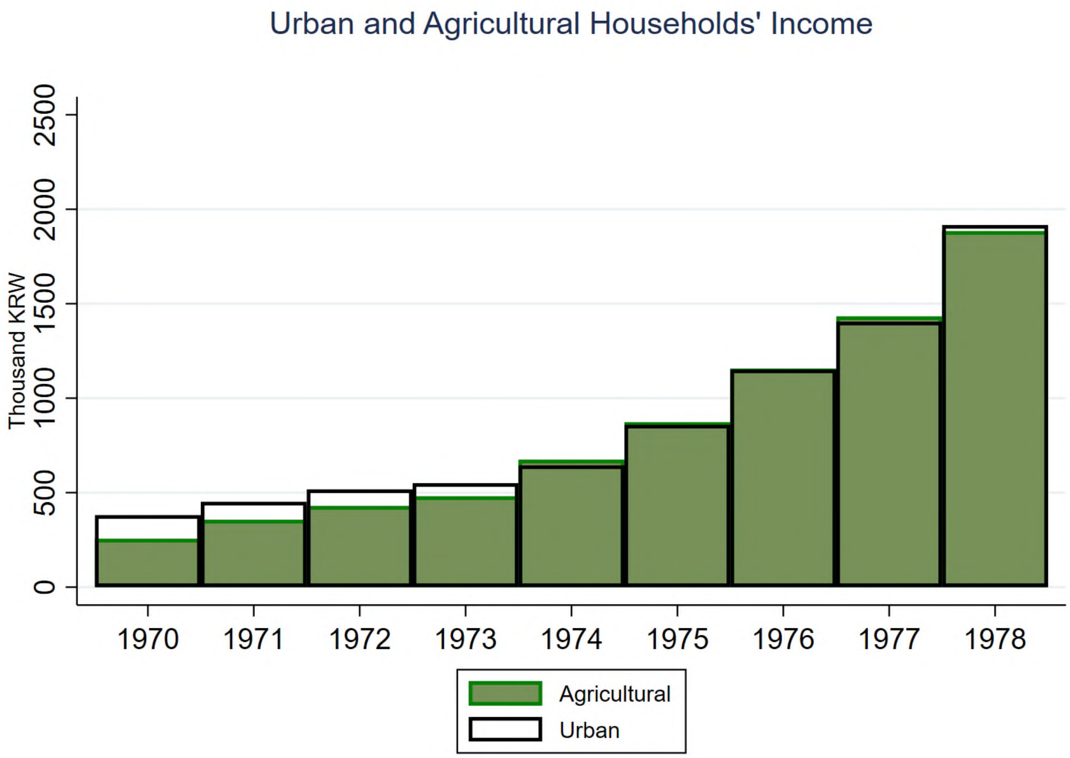


Figure A.4: Income Gap Between Urban and Agricultural Households in 1970s



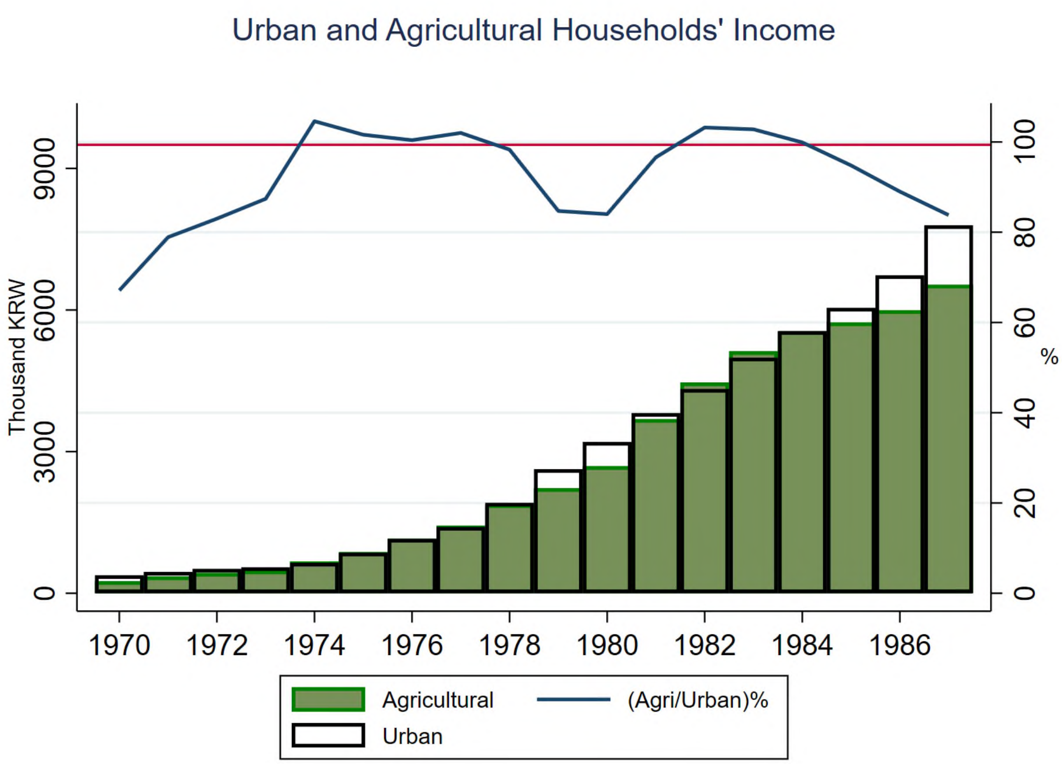


Figure A.5: CDD Budget in *Namhae* (KRW)

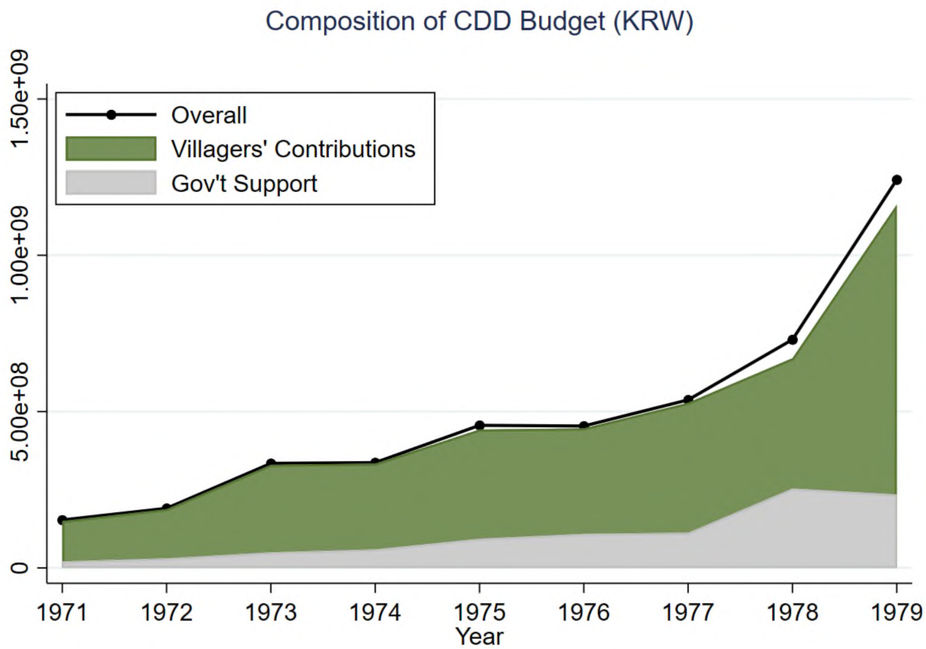


Figure A.6: An Example Page of OCR Result

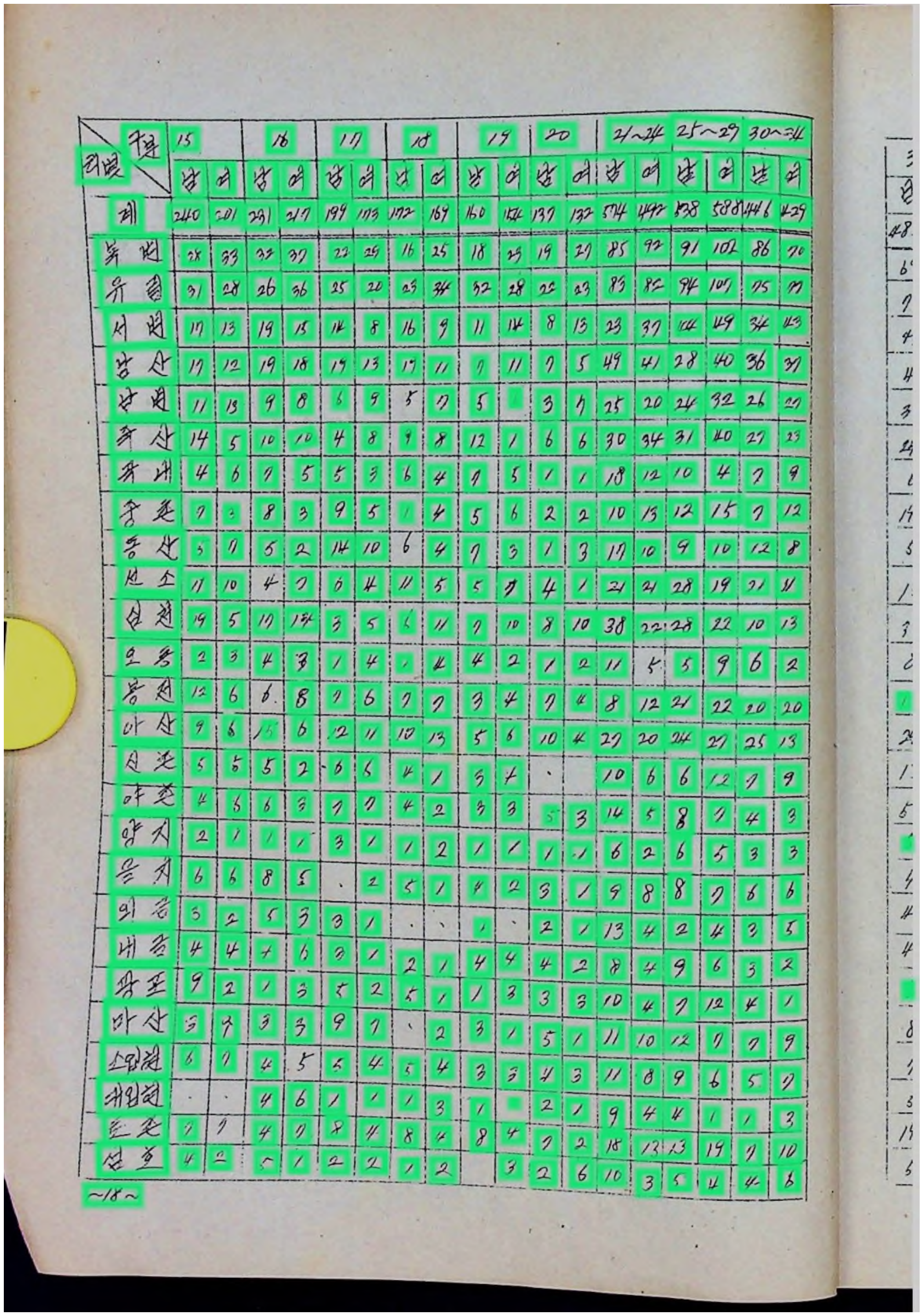
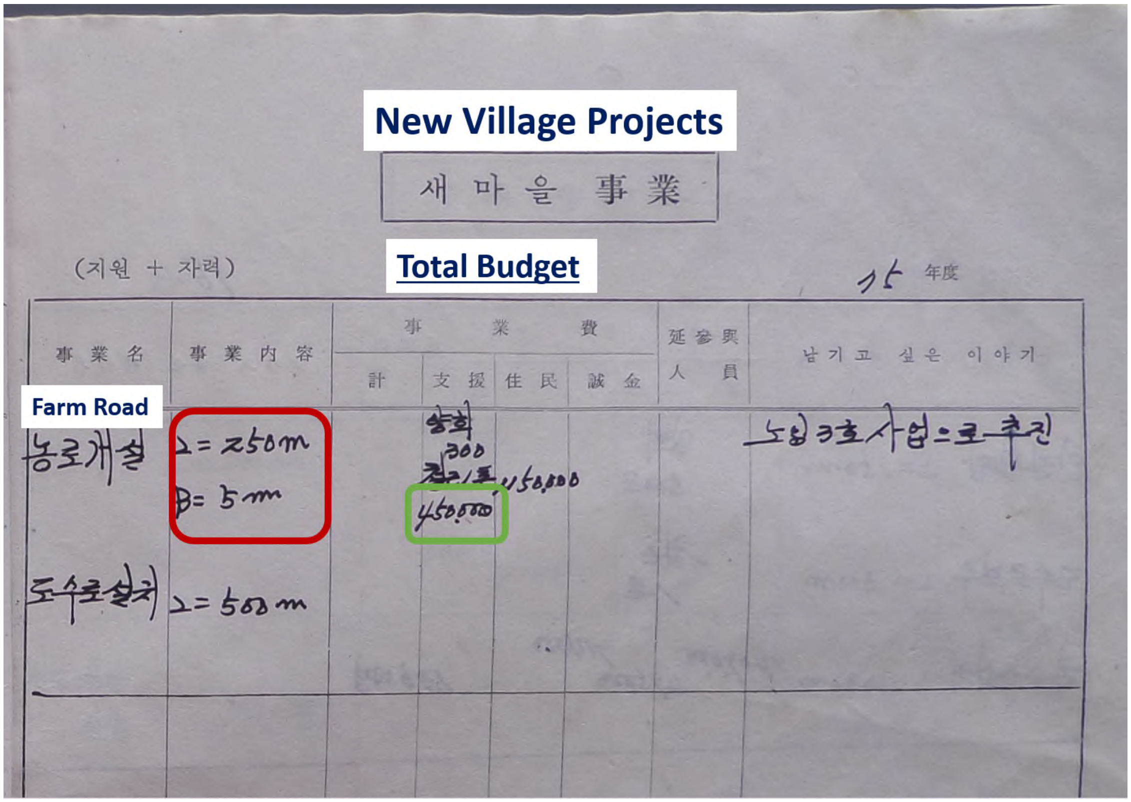


Figure A.7: *History of Our Village*

Figure A.8: Evidence of Validity of *History of Our Village*



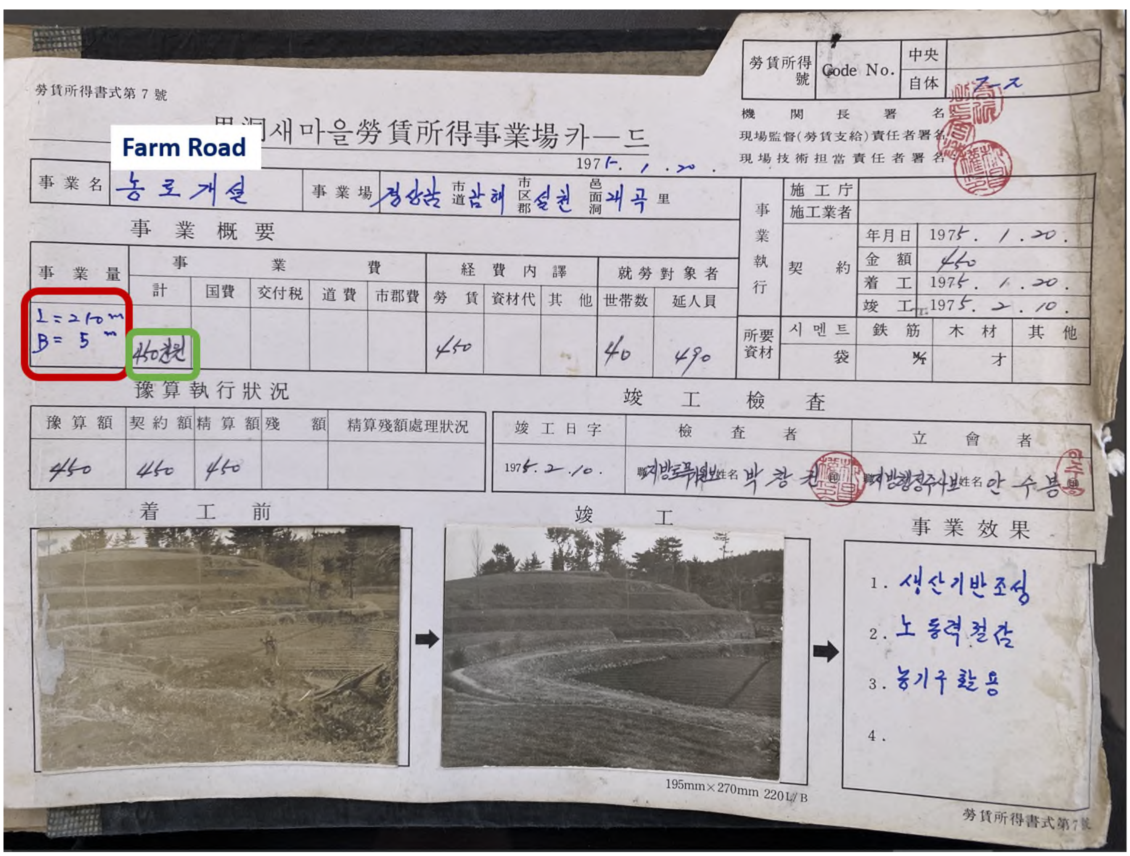


Figure A.9: An Example Page of the 1930 Census

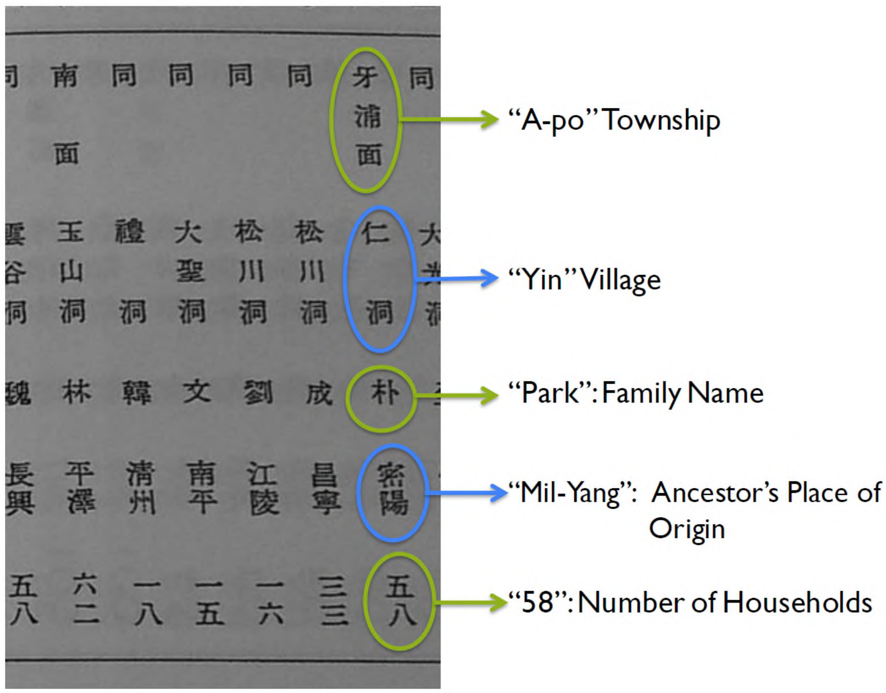


Figure A.10: Evolution of Rice Yield in Namhae County

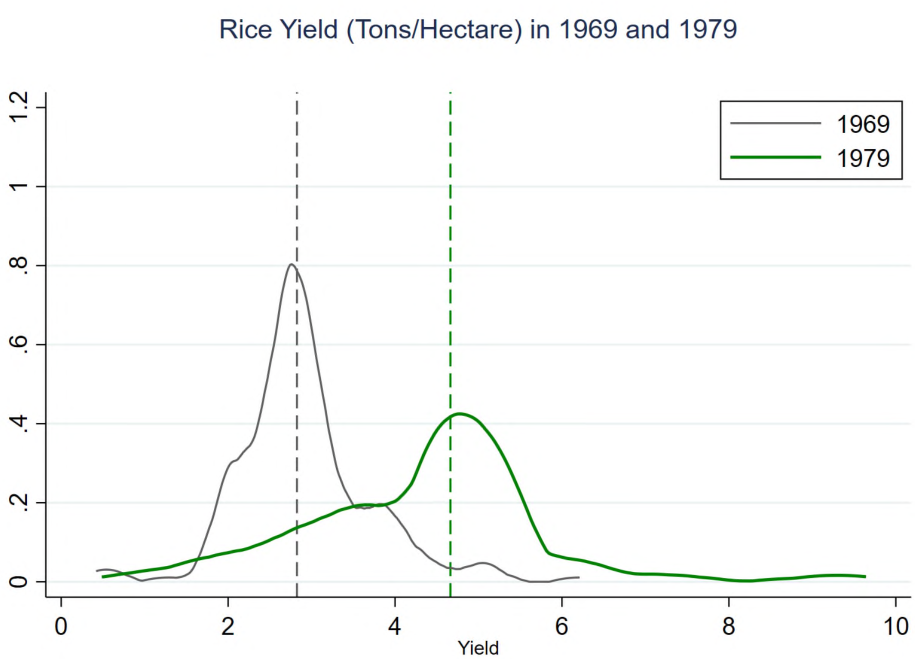


Figure A.11: The Effect of Homogeneity and the Bridge

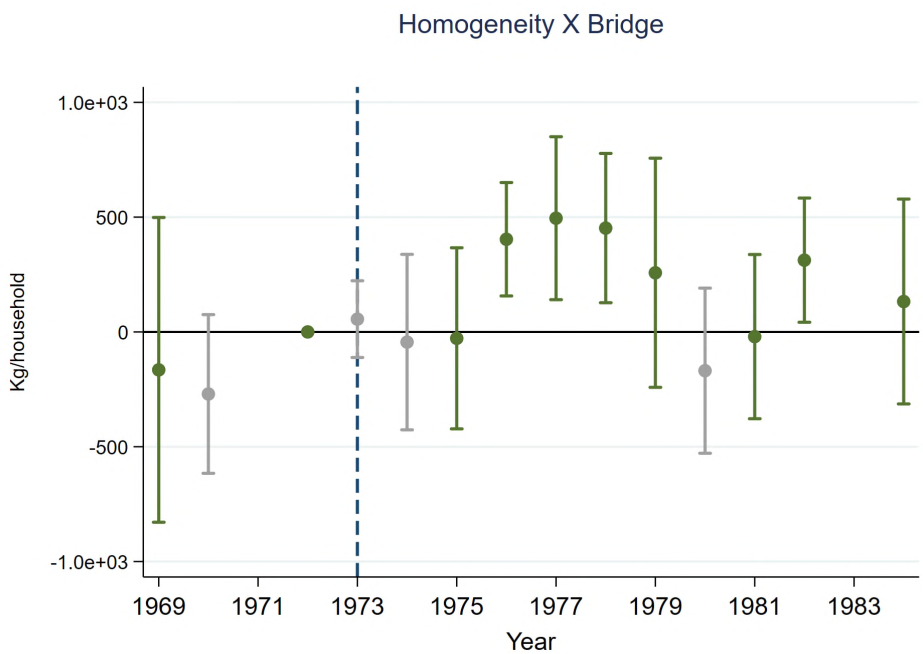


Figure A.12: The Effect of Homogeneity and the Bridge using Rice Yield

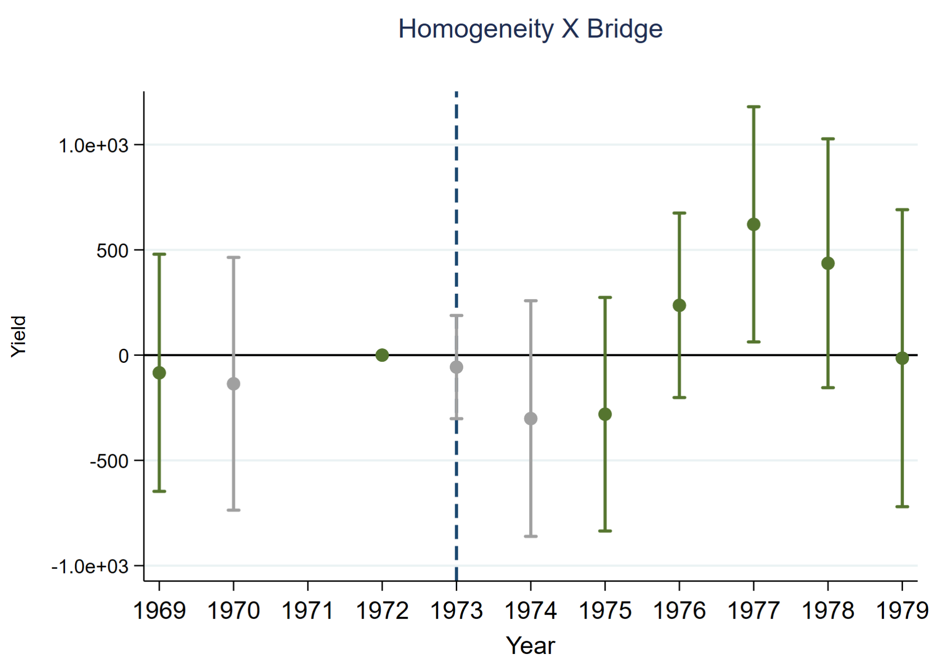


Figure A.13: Distances within *Namhae*

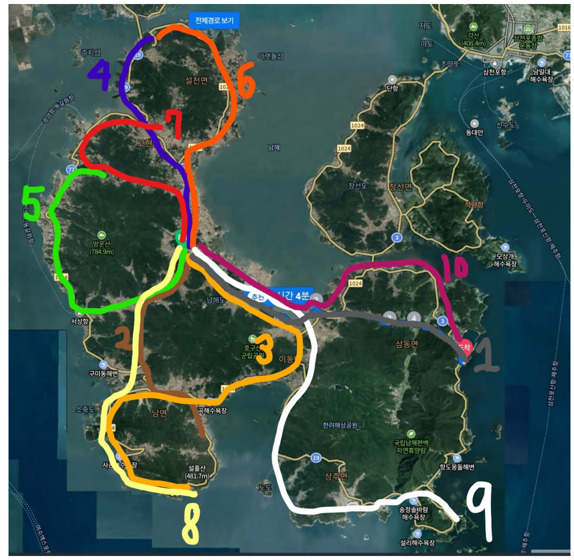
8. 주요이정표

남해							
5,1	초음						
8,2	3,1	이동					
11,4	6,3	3,2	영지				
14,0	8,9	5,8	2,6	갈곡			
16,1	11,0	7,9	4,7	2,1	봉화		
17,2	12,1	9,0	5,8	3.2	1,1	동천	
21,1	16,0	12,9	9,7	7,1	5,0	3,9	은점

남해											
8.2	이동										
12,7	4 5	화계									
13,5	5,3	0.8	용소								
17.2	9,0	4.5	3,7	두곡							
17,7	9.5	5,0	4,2	0,5	비로정						
19,2	11,0	5.5	4,7	1.0	0,5	당항					
20,3	12,1	7.6	6,8	3,1	2.6	2.1	양지				
23.5	15,3	10,8	10,0	6,3	4.8	4,3	2.2	평산			
24,6	16,4	11,9	11,1	7,4	6,9	6,4	4.3	2,1	유구		
28,8	20,6	16,1	15.3	11.6	11,1	10,6	8.5	6,3	4,2	항촌	
33,2	25,0	20,5	19,7	16.0	15,5	15.0	12,9	10,7	8.6	4,4	가천

남해	11,6	13.0	15.0	15.9	17,7
	양지	1.4	3.4	4.3	6,1
		당항	2,0	2.9	4,7
			석교	0.9	2,7
				숙호	1.8
					홍현

(a) Distance Matrices



(b) A Sketch of Road Network

Table A.11: Summary Statistics by Clan Homogeneity and Access to Bridge

	(1) $(H_i = 1, B_i = 1)$	(2) $(H_i = 1, B_i = 0)$	(3) $(H_i = 0, B_i = 1)$	(4) $(H_i = 0, B_i = 0)$	(5) All Villages
Baseline Characteristics (1972)					
Num. of villages	84	14	85	12	195
Population	570.5	724.1	684.1	836.8	647.4
Households	99.5	115.6	119.7	218.2	116.8
Agri. Households	89.9	98.9	87.7	105.7	90.6
Share modified roofs	0.31	0.31	0.35	0.37	0.33
Share of villages with schools	0.17	0.14	0.28	0.25	0.22
Market Access (1972)					
Distance to the bridge (km)	22.7	N/A	25.5	N/A	24.1
Share of car accessible villages	0.75	0.93	0.79	0.33	0.76
Share of highway accessible villages	0.23	0	0.24	0	0.20
Share of state road accessible villages	0.15	0.5	0.16	0.33	0.19
Share of county road accessible villages	0.55	0.57	0.34	0.50	0.45
Agriculture (1969, except for tillers)					
Rice output per agri. household (kg/hh)	924.8	833.5	773.7	957.9	854.3
Rice yield (kg/ha)	1906.5	1278.1	1609.5	1427.8	1705.4
Cultivated areas (ha)	43.4	49.5	42.3	64.9	45.6
Num. of tillers per household (1976-1979)	0.048	0.023	0.042	0.031	0.041
Village Projects (1970-1979)					
Avg. budget per year (USD)	$3,998	$3,761	$6,513	$6,641	$5,143
Share of budget on infrastructure	0.42	0.27	0.42	0.28	0.40
Distance to Administrative Units					
Distance to the closest admin office (km)	4.4	4.8	5.9	3.5	5.0
Distance to the county office (km)	9.4	N/A	13.0	N/A	11.3

Note: H_i is an indicator variable that equals 1 when the Herfindahl index of clan concentration is greater than the median. B_i is an indicator variable for villages connected by the construction of the bridge in 1973. This table presents the means and standard deviations in parenthesis. Column (1) corresponds to the villages that are both homogeneous and have access to the bridge. Column (2) shows summary statistics for homogeneous villages without access to the bridge; Column (3) for heterogeneous villages with access; Column (4) heterogeneous villages without access. Column (5) shows the statistics for all villages.

Table A.12: The DDD Estimates of Clan Homogeneity on Agricultural Productivity

	Dependent variable: Rice Output/Agri. Household		
	(1)	(2)	(3)
	$1[H_i > med]$	$1[\text{TOPSHARE}_i > med]$	$1[\text{CLANSHARE}_i > med]$
Constant	1,031.32***	1,014.55***	1,029.80***
	(94.48)	(101.89)	(72.91)
Homogeneity	-157.95	-118.62	-217.20*
	(119.18)	(124.33)	(116.16)
Bridge	-299.85***	-277.86**	-279.84***
	(103.02)	(109.96)	(85.38)
Homogeneity × Bridge	340.92***	294.29**	240.84**
	(130.80)	(135.58)	(121.93)
post	392.96***	384.34***	358.10***
	(106.17)	(115.17)	(84.51)
Homogeneity × post	-229.93*	-199.39	-231.23**
	(122.09)	(129.94)	(111.58)
Bridge × post	-42.00	-40.56	14.25
	(113.41)	(121.84)	(94.66)
Homogeneity × Bridge × post	**285.01****	**268.33***	**288.46****
	(130.90)	**(138.28)**	**(121.87)**
Observations	1,338	1,338	1,338
R-squared	0.13	0.12	0.11

Note: This table presents the DDD estimates for the three measures of clan homogeneity (H_i, $TOPSHARE_i$, and $CLANSHARE_i$). Standard errors clustered at the 1930 village-level

A.2 Estimating the Distances to the Bridge

In this subsection, I describe how I estimate the distances from villages to the bridge using distance matrices published in *Namhae* County Statistical Yearbooks. Figure A.13a shows the distance matrices for three of the 10 local roads from the 1970 Statistical Yearbook. Distance matrix provides pairwise distances between destinations along a route. Most of the stopovers are village centers, but when a road does not pass through a village, the name of region is provided. I have roughly sketched the road network based on the distance matrices in Figure A.13b.

I calculate the shortest route between a village and the bridge by considering all possible travel routes and taking the minimum.

(1) If a village has a direct connection to the bridge by a road (i.e., without switching to another road), then I take this direct route as the shortest travel distance. There are 18 such villages.

(2) Most villages need more than one road to get to the bridge. I identify the villages that take the above 18 villages as destinations, then add the two distances (adding the travel distance from a village to one of the 18 directly-connected villages, and the distance from the directly-connected villages to the bridge). There are 62 such villages that need a stopover.

Out of the 170 villages in the connected island, this process identifies accurate distances for 80 villages (47%). I assume that the villages within the same "legal" village boundaries have similar distances to the bridge. Legal village boundaries are first drawn during the colonial period, and are still widely used in legal documents. However, as population grew, legal village boundaries were further divided into "administrative" villages for administrative convenience. Such divisions are not arbitrary—they are drawn based on the locations of communities within a legal village. In general, there are two or three "administrative" villages within a "legal" village boundary. Villages within the same legal unit are typically within 1 kilometer apart. Assuming that villages within the same legal village boundary share similar routes to the bridge gives me the distances for 60 more villages.

For the remaining 30 villages without the distances, I use Google Maps to find contemporary walking distances to the closest village, for which I know the 1970's distance to the bridge. For instance, if village A's distance is missing, but the adjacent village B's distance is known from the 1970 data, then I calculate village A's distance by assuming village B as a stopover.

Table A.13: Summary Statistics: Distance to the Bridge (Km)

	N	Mean	SD	Min	Max
Distance	170	24.1	11.6	0	51.9
By Township					
Namhae	26	18.3	1.53	15.4	20.8
Leedong	26	29.2	7.17	21.2	43.1
Samdong	31	37.5	7.9	26.1	51.9
Nam	24	32.1	3.82	26.4	41.7
Seo	23	24.6	5.7	15.7	34.9
Kohyun	21	11.1	4.9	5.1	24
Seolchun	19	6.5	3.76	0.1	12.2
*Changsun**	0	0	0	0	0

Note: Changsun is the control island with 23 villages.

A.3 Data: Further Documentation

Types of Projects in *History of Our Village*: *History of Our Village* has detailed records of $5,271$ village development projects from 120 villages in *Namhae* county between 1971 and 1982. *Namhae* county had 195 villages in 1971, so this village history book covers 62% of the villages (107 of 169 in the treatment group (63%), and 13 of 26 in the control group (50%)). Of the $5,271$ projects, there were 266 unique projects.

Table A.14 shows 15 most frequent projects by the treatment status. This is in line with the list of major projects from *The Ten Years of New Village Movement* that was published by the Ministry of Home Affairs in 1980 to commemorate its achievements. Most of the projects were small-scale infrastructure projects that could be implemented and managed by village residents with the support from local governments.

Table A.14: Most Frequent Projects

	Total	Treatment Villages	Control Villages
Total (Entire data set)	5,271	4,420	851
15 Most Frequent Projects			
Housing Improvement (roof amelioration)	322	261	61
Housing Improvement (fence)	308	229	79
Construction of Farm Roads	290	262	28
Expansion of Village Roads	210	164	46
Small Bridges	208	182	26
Modernizing Kitchens	178	138	40
Storage Houses	141	126	15
Installing Water Ways	137	118	19
Small River Arrangements	137	104	33
Public Restrooms	137	96	41
Village Halls	130	121	9
Farm Roads Repairs	126	109	17
Housing Improvement (roof painting)	124	96	28
Sewage Systems	108	102	6
Pavement of Village Roads	88	82	6

Categorization of Types of Projects: I categorize 238 unique projects into 15 broad categories. Table A.15 show the details of classification. Note that the broad categories are not mutually exclusive. Some projects—such as *construction of water ways*, are included in both *village water management* and *agricultural irrigation*, as those two broad categories themselves are not mutually exclusive.

Table A.15: Categorization of Village Projects

Broad Categories	Detail (*# of the unique projects under each broad category*)
Village Roads	Construction/Expansion/Repair/Pavement/Widening of village/farm/entry/other roads (23)
Village Bridges	Construction/Expansion/Repair of village bridges (3)
Ports	Construction/Expansion/Repair of village ports (3)
Village Water Management	Construction/Expansion/Repair of portable water supply facilities; Construction/Repair of water ways, public wells, and reservoir banks; Establishment/Repair/Covering of sewage system (20)
Village Stream Management	Construction/Repair of breakwaters, weirs (low head dams), reservoir banks, dams; Cleaning/Deepening of river bottom (22)
Roof Improvements	Modernizing/Painting roofs (3)
House Improvements	Renovating houses; Remodeling fences/walls/kitchens (5)
Electricity	Electrification (2)
Miscellaneous Facilities	Public baths, Laundry places, Children's parks, Senior centers, Street lights, Restrooms, Salons, Libraries, etc. (55)
Village Centers	Construction/Expansion/Repair of village centers (3)
Village Warehouse	Construction/Repair of village warehouses or silos (3)
Agriculture	Land reclamation, Vinyl houses, Orchards, Farming machines, Introducing high-yield varieties, Farmland irrigation/arrangement, Construction/Repair of water ways and reservoir banks, Construction/Repair of village warehouses or silos, etc. (57)
Agricultural Irrigation	Irrigation/Arrangement of farmland, Stream channel straightening, Construction/Repair of water ways and reservoir banks, etc. (29)
Foray into New Businesses	Raising diary cattle, Stockbreeding, Oyster farming, Other shellfish farming, Nursery gardens, etc. (31)
All Others	Activities of Women's Association, Training Sessions, Cleaning village roads, Afforestation (32)