Road Quality and Local Economic Development: Evidence from Indonesia's Highways*

Paul J. Gertler[†]
University of California, Berkeley

Marco Gonzalez-Navarro[‡]
University of California, Berkeley

Tadeja Gračner[§]
RAND Corporation

Alexander D. Rothenberg[¶]
Syracuse University

December 2021

PRELIMINARY

Abstract

This paper provides evidence on the crucial role that highway quality has on economic development. We propose a new instrument for road quality driven by features of Indonesia's two-step budgeting process for allocating funding for road improvements to different road authorities. We first show that higher road quality leads to job creation in the manufacturing sector mostly by expanding the number of new firms. We then show that this is reflected in household consumption and income. Third, we show evidence of an occupational shift from the informal sector of employment into manufacturing and improved profits for those who stay in agriculture. This is consistent with highway quality being an important productive amenity for agriculture and industry in developing countries.

JEL Classifications: J43, O18, R23, R42

Keywords: transport infrastructure, road maintenance, economic development

*The authors thank seminar participants at UC Berkeley, Millennium Challenge Corporation, the RAND Corporation, University of Toronto, and University of Ottawa for helpful comments and suggestions. This work was generously supported by a USAID Economic Research Partnership Grant.

[†]Haas School of Business, MC #1900 Berkeley, California 94720. Email: gertler@haas.berkeley.edu.

[‡]Ag. Economics, MC #3310, Berkeley, California 94720. Email: marcog@berkeley.edu.

[§]1200 South Hayes St., Arlington, VA 22202-5050. Email: tgracner@rand.org.

[¶]426 Eggers Hall, Syracuse NY 13244-1020. Email: adrothen@maxwell.syr.edu.

1 Introduction

One of the major functions of government is the maintenance and improvement of existing road networks, e.g. repaving, and is often justified as a public good investment to stimulate economic activity and create jobs. Although most public expenditures on roads are designated for road maintenance, the paucity of credible evidence on the impacts these investments have on local economic development is surprising.[1] In this paper, we aim to fill this knowledge gap by analyzing how changes in national and provincial highway road quality impact local economic development outcomes in Indonesia.

Our empirical analysis uses an unusually long and comprehensive administrative database on road quality. From 1990 to 2007, a period of rapid economic transformation, Indonesia's highway authorities measured the roughness of each road segment covering three of the country's most important islands: Java, Sumatra, and Sulawesi. We combine these spatio-temporal measures of road quality with several high-quality datasets capturing economic outcomes, including: (1) the Indonesia Family Life Survey (IFLS), a nationally representative household panel database; (2) the Industrial Survey (SI), an annual census of large manufacturing firms; and (3) population census data.

Causal estimates of the effects of road improvements are difficult to obtain because of the endogeneity of maintenance decisions. If planners target particular corridors for road improvements based on economic or political characteristics, this generates selection bias (Blimpo et al., 2013; Burgess et al., 2015; Asher and Novosad, 2020). We overcome this challenge with a novel instrumental variables strategy that takes advantage of Indonesia's centralized fiscal organization and budgeting process.

In Indonesia, independent road authorities corresponding to different government tiers (e.g. national, provincial, and district) make investment decisions based on a two-stage budgeting process. In the first stage, the central government sets an annual total budget for maintenance and this common pool is subsequently allocated to units based on publicly available formulas. We use a large vector of observable characteristics at the district-level to proxy for these formulae that govern road maintenance decisions. Then, in the second stage, different road authorities use the funds allocated to them to upgrade their choice of roads. This implies that endogeneity of road investments is limited to second stage decision making since a particular road authority's budget only depends on dimensions such national tax revenues and national investment formulas that are not manipulable.

This two-step budgeting process forms the basis for two different sets of instruments: (1) a district's own baseline characteristics interacted with the total provincial budget for road investments; and (2) the sum of other districts' characteristics interacted with the total provincial budget for road investments. Because we have a large vector of characteristics from which to construct instruments, we use lasso techniques to choose instruments that give us the best fit in the first stage, following Belloni et al. (2012).

Using this identification strategy, we first verify that for villages in the IFLS, improvements in district-level road quality reduce travel times to nearby district and provincial capitals, as would be expected. These travel time reductions are observed both in IFLS survey responses and in measures of travel times derived from the engineering data.

Next, with district-level panel specifications based on the manufacturing census, we find that the total value added, output, and employment of manufacturing firms in the district increases in response

[1] Engineering models from the World Bank (1994) estimate that the returns to road maintenance are twice as high as those for network expansions.

to improvements in road quality. In addition, because the employment response is lower than the impacts on output, output per worker increases. We exploit the fact that we have firm-level panel data to demonstrate that these aggregate effects are driven by new firms, and not by changes in output or output per worker for incumbent firms.

We then turn to the Indonesia Family Life Survey, which allows us to better understand the sources of the increased labor we observe in manufacturing. These data show that when road quality improves, we observe an occupational shift away from informal employment (including agriculture) and towards manufacturing jobs. The IFLS also allows us to understand road quality's impacts on small scale agriculture. In particular, we find that road quality increases the returns to agriculture (higher profits for those who stay in agriculture). Despite the fact that we neither observe extensive margin effects on employment nor intensive margin effects on hours worked, both the industrial employment expansion as well as the more productive farming imply that total household earnings are clearly larger, and this is corroborated when using measures of consumption per capita.

Finally, using cross sectional district level census data, we provide suggestive evidence that better local roads lead to small amounts of in-migration of households. Road quality improvements reduce the price of perishable food products, and we also find suggestive evidence that these road quality improvements increase land values.

This paper contributes to a sizeable literature evaluating the impact of transport infrastructure improvements in developing countries. Several important papers study the effects of newly created surface links that expand transportation networks, including China's national trunk roads, (Banerjee et al., 2012; Faber, 2014), India's Golden Quadrilateral Project (GQ) (Ghani et al., 2016), new railways in colonial India (Donaldson, 2018), or new highways in Brazil (Morten and Oliveria, 2018; Bird and Straub, 2020). Our work is different in that it focuses on quality improvements to existing roads, which are relatively understudied. Such projects may be less expensive and more politically feasible, potentially resulting in a higher benefit-cost ratio.

Our work is also related to a large body of evidence on the impacts of transport improvements to rural areas, which often involves upgrades to existing unpaved roads (e.g. Aggarwal, 2018; Gollin and Rogerson, 2014; Khandker et al., 2009; Valdivia, 2011; Khandker and Koolwal, 2011). Like our work, Asher and Novosad (2020) find that new rural roads due to the Village Road Program in India lead to a transition of workers out of agriculture, but there were no impacts on income or migration, unlike in our work. Our research identifies the impacts of improvements to national and provincial highways, which are far more important for the movement of goods and services, potentially explaining differences in the effects. Another advantage of our work is that it benefits from a continuous road quality measure, instead of binary treatment that is often used in the literature.

Finally, a large body of work studies the impact of transport infrastructure improvements through the lens of trade theory. For example, Donaldson and Hornbeck (2016), Asturias et al. (2019), and Storeygard (2016) find large, positive effects of transportation infrastructure on aggregate welfare and income, but for these studies, the impacts largely came about due to a reduction in transport costs and an increase in trade volumes. We complement that work by viewing transport infrastructure improvements as a productive amenity and show that Indonesia's road improvements attracted firms and workers, lead to a transition of labor out of agriculture, and increased housing prices. This approach complements work in urban economics on highway improvements that has mostly focused on the U.S. interstate highways system. For example, Duranton and Turner (2012) investigate city growth effects, Michaels (2008) analyzes skill premia changes, while Baum-Snow (2007) documents suburbanization effects.

The rest of the paper proceeds as follows. Section 2 provides a theoretical discussion of how improvements to existing roads may affect local economic development outcomes, and section 3 describes the datasets used. Section 4 describes the historical and institutional background behind the evolution of road quality and discusses the identification strategy. Section 5 presents our results, and section 6 concludes.

2 Theoretical Framework

In this section, we describe a simple theoretical framework for understanding how improvements to road quality may impact local economic development. There is a collection of M discrete locations (e.g. districts), indexed by $m = 1, ..., M$. Each location is endowed with a $(K \times 1)$ vector of amenities, $\theta_m = (\theta_1, \theta_2, ..., \theta_K)'$. Local road quality, θ_r, is one element of this vector θ.

As in Jacoby (2000), road quality is valued as a productive amenity by agricultural producers, because it enables them to reduce input costs and more efficiently get their products to market. Manufacturing firms value road quality for similar reasons. Workers also positively value road quality as a consumer amenity, both because of potentially lower consumer prices but also because road quality improvements may reduce commuting times and vehicle maintenance costs (Redding and Turner, 2015).

We also emphasize the dual-economy nature of many developing economies, where there is a high wage, productive formal sector, and a low-wage, unproductive informal sector (LEWIS, 1954; Temple, 2005; La Porta and Shleifer, 2014). In the formal sector dominated by manufacturing in our context, firms produce a composite good, X, under a constant returns-to-scale production environment. This good is freely traded over space, and without loss of generality, we can normalize output prices to 1, so that this good is the numeraire of the model. Firms use capital, labor, and land for production.

For some values of road quality (i.e. when $\theta_r < \overline{\theta}$), no formal manufacturing firms will be able to produce with *any* amounts of capital or labor. This amenity threshold introduces a non-concavity in an otherwise standard production function. In such locations, formal sector employment opportunities are nonexistent, and workers instead supply their labor to the informal sector. Agriculture represents a large source of informal sector employment (e.g. Singh et al., 1986; Benjamin, 1992; Bardhan and Udry, 1999).

Households make choices about where to live, and conditional on residential choices, they choose a sector in which to work. When road quality improves, manufacturing firms are now able to produce in locations that were previously infeasible. Those firms move in, bringing with them new, higher wage formal sector jobs that were not there before. These new employment opportunities encourage workers to switch sectors, moving out of informal employment and into formal sector jobs. This increases workers' total earnings and consumption, because those workers now have access to formal employment opportunities and earn higher wages.

In the absence of migration costs, workers are perfectly mobile between locations. Because workers consume housing, which has an upward-sloping supply curve, when more people move to a community, housing prices increase. In a spatial equilibrium model, wages and housing prices will adjust to make workers indifferent between locations, as is standard in labor and urban economics (e.g. Rosen, 1979; Roback, 1982). In such a model, the new employment opportunities should initially raise wages and encourage in-migration of workers, increasing housing values. However, if internal migration is costly (e.g. Bryan and Morten, 2019), the new manufacturing firms may be able to bring jobs to a community, raise wages, and lead to positive welfare benefits for the affected communities that are not completely bid-away through housing price effects.

In our empirical work below, we test each of the different aspects of this theoretical framework, in which road quality improvements are expected to: (1) lead to new manufacturing jobs; (2) encourage workers to switch sectors; (3) increase total earnings and consumption; (4) encourage greater in-migration; and (5) increase housing prices and land values.

3 Data

To study the effect of road maintenance on local economic development outcomes in Indonesia, we combine several high-quality data sources. These include road quality measures from engineering surveys, household survey data, manufacturing data, population census data, and geospatial datasets. We briefly describe each of these data sources in turn, and more details can be found in Appendix B.

Road Quality Data. Every year, to help manage its surface transportation network, Indonesia's Department of Public Works (*Departemen Pekerjaan Umum* or DPU) conducts a high resolution data collection effort to monitor pavement quality. Surveyors collect information on every segment of national and provincial highways, with measures that include the surface type, width, and road roughness. Our data includes this information for all provincial and national roads in Indonesia for the years 1990-2007, and contains more than 1.2 million kilometer-post-interval-year observations. The road quality data were also merged to maps of road networks to provides us with an annual spatial panel of road quality measures.

To create a summary measure of road quality, we use the international road roughness index (IRI), a widely accepted measure of road quality in civil engineering that was developed by the World Bank in the 1980s.[2] The IRI is defined as the ratio of a vehicle's accumulated suspension motion (in meters), divided by the distance traveled by that vehicle (in kilometers) during measurement. All else equal, when driving on inferior roads or when faced with potholes and ragged pavement, drivers decrease speeds and prolong their travel time. Rough surfaces also increase accidents, add to maintenance costs, and require greater fuel consumption. Consequently, road roughness is directly related to local productive and consumer amenities both in terms of travel times, vehicle maintenance costs, and increased accidents (Bock et al., 2021).[3]

Let $\mathcal{R}_d$ denote the set of national and provincial road segments in district d, and let d_r denote the length of road segment r. We measure the average road quality in district d as follows:

$$\text{Road Quality}_{dt} = (-1) \times \frac{\sum_{r \in \mathcal{R}_d} d_r \text{IRI}_{rdt}}{\sum_{r \in \mathcal{R}_d} d_r} \tag{1}$$

[2]See Appendix Section B.1.2 for more details on the IRI.

[3]Fuel consumption and labor costs account for more than 50 percent of vehicle operating costs in Indonesia (Asia Foundation, 2008).

where IRI_{rdt} denotes the roughness of road section r in district d at time t. Hence, Road Quality$_{dt}$ is the negative of a distance-weighted average of road roughness for all roads in that district. Importantly, this average is taken over all national and provincial roads located in the district, and different districts have different shares of each of these types of roads.

Figure 1 shows substantial variation in road quality over time, observed as a significant leftward shift in the distribution of average road roughness across districts between 1990 and 2000. Similarly, Figure 2 documents substantial spatial variation in road improvements over time on the Island of Sumatra. Nearly 84 percent of Sumatra's network of national and provincial highways were unpaved in 1990, but this figure fell to 46 percent only a decade later. Similar trends apply to Java's highway network.[4] Finally, Figure 3 shows that the distribution of road quality substantially narrowed between 1990 and 2007, suggesting that the maintenance projects were targeted at improving quality at the low end of the distribution.

Indonesia Family Life Survey (IFLS). Our main source of data for individual and household-level labor market and consumption outcomes is the Indonesia Family Life Survey (IFLS). The IFLS is a national longitudinal survey, representative of 83 percent of Indonesia's population, and it tracks more than 30,000 individuals in 5 waves over a 19 year period. These individuals are observed in more than 300 villages (*desa*), which are located in 13 of Indonesia's 27 provinces and in over 200 districts (*kabupaten*). Although the most recent survey wave was conducted in 2015, we only use data from the first four waves in 1993, 1997, 2000, and 2007 to match the timing of the road quality data. Figure 4 shows the locations of IFLS villages used throughout our analysis. The IFLS is notable for its low attrition rate, as more than 87 percent of the original households were tracked through the first four waves of the survey. Consequently, these panel data allow us to track the same households and individuals facing different road infrastructure conditions over almost twenty years.

Census of Manufacturing Firms. Our primary data source for firm-level outcomes is the Annual Census of Manufacturing Establishments (*Survei Tahunan Perusahaan Industri Pengolahan*, or SI), collected by Indonesia's central statistical agency, (*Badan Pusat Statistik* or BPS). The SI is an annual census of manufacturing plants with more than 20 employees and contains detailed information on plants' cost variables, employment size, and measures of value added. One advantage of SI data is that they contain firm-level identifiers, allowing us to track changes in firm-level outcomes over time. The data also contain information on plants' starting dates, locations at the district level, as well as firm-level outcomes, such as employment and wage rates, value added, and output.[5]

Population Census. To study population and migration responses to changes in road quality, we combine the above datasets with data from the 2000 Population census. These data collect individuals' birth districts and other socio-demographic characteristics. We also supplement the 2000 Population census data with data from the 1971, 1980, and 1990 censuses from the Integrated Public Use Microdata Series (IPUMS).

[4] See Appendix Figure A.1 and Appendix Figure A.2.

[5] New firms are counted when they appear in the dataset having never appeared before. Also, for the purpose of our analysis, we dropped all firms coded as state-owned enterprises (less than 3 percent of all firm-year observations). Throughout the discussion, we use plants and firms interchangeably since less than 5% of plants in the dataset are operated by multi-plant firms (Blalock and Gertler, 2008).

Geospatial Data on Administrative Boundaries and Topography. Our analysis also relies on administrative boundary shapefiles that identify district borders. These data are created by Indonesia's national statistical agency, *Badan Pusat Statistik* (BPS). We use these boundaries in combination with data from the *Harmonized World Soil Database (HWSD)* to construct several measures of basic topographic characteristics (e.g., area, ruggedness, slope, and elevation).

Policy Variables. Finally, we also use data from the Indonesia Database for Policy and Economic Research (INDO-DAPOER), maintained by the World Bank, to construct time-varying economic and social indicators at the district-year level. These indicators span four main categories: fiscal, economic, social, and demographic.

4 Empirical Strategy

Our objective is to identify the causal effect of road quality on local economic development outcomes. To do so we estimate the following regression model:

$$y_{dt} = \alpha_d + \alpha_t + \beta \log \left(\text{Road Quality}\right)_{dt} + \mathbf{x}'_{dt}\theta + \varepsilon_{dt}, \tag{2}$$

where y_{dt} is the outcome of interest for district (or community) d at time t and $\log\left(\text{Road Quality}\right)_{dt}$ is defined in (1). The vector $\mathbf{x}_{dt}$ represents a set of time-varying controls, including non-oil district-level GDP and population (both in logs), and α_d and α_t represent district and year fixed effects. Panel data allow us to control for time-invariant unobservables that may be correlated with changes in road quality and outcomes, while year-fixed effects allow us to control for any national, year-specific changes that may drive both outcomes and road quality. If y_{dt} is measured in log-terms, the key parameter of interest, β, measures the elasticity of y with respect to road quality. In some specifications, we estimate a regression model similar to (2) at the individual and household level; in that case, we include corresponding control variables and individual or household fixed effects, respectively.

Causal estimates of the effects of road improvements are difficult to obtain because of maintenance decisions are endogeneous, and the sign of the selection bias is difficult to ascertain a-priori. On the one hand, if planners targeted faster growing areas with road improvements, this would create a positive selection bias in parameter estimates, even in specifications with district fixed effects. On the other hand, areas with faster economic growth will tend to experience more rapid road deterioration when roads are more intensively used, and this could create a negative selection bias in parameter estimates. Neither of these problems are completely solved with time and location fixed effects.

We can describe both concerns using a capital accumulation formulation for the evolution of road quality:

$$\text{Road Quality}_t = \text{Road Quality}_{t-1} \cdot (1-\delta) + \alpha I_t \tag{3}$$

where Road Quality_t measures road quality for a given segment at time t, δ denotes the deterioration rate, and I_t measures investment in that segment at time t. Instead of specifying a fixed per-period deterioration rate, suppose that greater usage causes road quality to deteriorate more rapidly:

$$\delta \equiv \delta_t = \phi_1 + \phi_2\,\text{Use}_t$$

Using this assumption, we can rewrite (3) as

$$\text{Road Quality}_t = \text{Road Quality}_{t-1} \cdot (1 - \phi_1 - \phi_2 \, \text{Use}_t) + \alpha I_t \tag{4}$$

Equations (3) and (4) show how estimates of the effects of changes in road quality may be confounded. Areas receiving improvement investment ($\uparrow I_t$) might have been selected by policymakers because of previous rapid deterioration or for economic growth or expected growth reasons. Second, if better roads increase local economic activity, this feedback may generate attenuation bias as roads may deteriorate faster due to their extensive use ($\uparrow \text{Use}_{rt}$).

To address these concerns, we use a novel instrumental variables strategy that takes advantage of Indonesia's centralized fiscal organization and budgeting process. Indonesia uses a two-stage budgeting process where budgets were allocated to local authorities that then decided how to allocate those budgets to specific road segments. While the second stage is clearly endogenous, we show in the next sub-section that the first stage, i.e. the budget allocations to local road authorities, is plausibly exogenous and is a good instrument for road quality. To construct this instrument, we take advantage of Indonesia's administrative decision-making process for road maintenance investments. In the following sub-sections, we describe this process and the instrumental variables we use.

4.1 Road Maintenance Financing and Allocation

During their rule, Dutch colonists built and maintained much of Indonesia's current road network.[6] After independence in 1945, roads were left to deteriorate until 1967, when Suharto assumed power. Road rehabilitation and improvement then became a top priority, and quantitative targets for improvement were included in many national five-year development plans (*Rencana Pembangunan Lima Tahun*, or *Repelita*). Spending on roads increased rapidly until the late 1970s, but it slowed in response the collapse of state oil revenues and remained stagnant during the 1980s. However, in the early 1990s, manufacturing began to grow rapidly and road rehabilitation and upgrading again became a priority.

During Repelita IV (1984-1989), the total budget for road improvements was $2.1 billion. This was increased by 84 percent in Repelita V (1989-1994), to a sum of $3.9 billion.[7] Transportation investments were the single largest item of the development budget during Repelita V, forming nearly 18 percent of total planned development expenditures. Almost all of the expenditures were allocated to improving the existing road network, especially upgrading dirt roads to asphalt. Although expenditures were planned to be kept at high levels during *Repelita* VI (1994-1999), the Asian financial crisis and its concurrent political upheaval resulted in less spending than originally intended. Road expenditures have experienced a slow recovery ever since (World Bank, 2012).

The Two Stage Budgeting Process: National and Provincial Roads. Maintenance and upgrading of national roads is primarily the responsiblity of the Directorate General of Highway Development (*Bina*

[6]Especially on Java, transport networks constructed by the Dutch were considered high quality by regional standards. By 1900, Java already had "a sophisticated agro-industrial economy integrated by overlapping networks of telegraphs, telephones, railways, and narrow-gauge tramways and good roads. Nowhere in Southeast Asia could boast better infrastructure. Elsewhere in East Asia, only Japan could compare" (Dick, 2000).

[7]These numbers, expressed in constant 2000 U.S. dollars, were taken from various planning documents describing Indonesia's five year development plans.

Marga) at the Ministry of Public Works and Housing (*Kementerian Pekerjaan Umum dan Perumahan Rakyat*, or Kemen PUPR). Under Suharto, the Ministry of Public Works carried out projects to upgrade national roads locally through its regional branch offices (*Kantor Wilayah*, or *Kanwil*) (Leigland, 1993). Funding for national road improvement projects came primarily through the central government's budget (*Anggaran Pendapatan dan Belanja Daerah*, or APBD) and was executed by direct ministry spending, typically on a project-by-project basis through *Daftar Isian Proyek* (DIP) funds.

In contrast, local public works agencies (*Dinas Pekerajaan Umum*, or *Dinas*) are responsible for maintaining and upgrading provincial highways. Provincial *Dinas* are answerable to the governors, operate with the support of local administrative agencies, and are funded through provincial budgets (Leigland, 1993; Lewis, 2017). Under Suharto, and especially during *Repilita V*, an important source of funding for provincial road improvements was *INPRES Jalan Propinsi*.

INPRES Jalan Propinsi was a central-to-provincial grant program, which began in 1979/1980, to fund the development and maintainance of provincial roads. Provinces were allocated *INPRES Jalan Propinsi* grants for improving roads from based on formulas, where allocation critieria included road condition, road length, density of roads, and per-unit construction prices (Shah et al., 1994). During *Repilita V*, funding for *INPRES Jalan Propinsi* increased substantially, from Rp 70 billion in FY 1989/1990 to Rp 348 billion by FY 1992/1993, a nearly 4 fold increase (Booth, 2003).

Under Suharto's *INPRES Jalan Propinsi* program, different provincial public works *Dinas* identified projects and submitted proposals to the central government. From the list of requests, projects were chosen and funds were released for approved projects after a "presidential instruction" was issued. The fund allocation process was top-down, with decisions made at the national level regarding total grant allocations, and total project budgets imposed from above (Crane, 1995). Hence, provincial road investment decisions can be thought of as following a two-stage budgeting procedure (Deaton and Muellbauer, 1980), where the central government sets the budget for each sub-national unit, and those budgets are independently allocates for spending.

After the fall of Suharto, the INPRES system of central to regional transfers was replaced by general allocation fund grants (*Dana Alokasi Umum*, or DAU) and special allocation fund grants (*Dana Alokasi Khusus*, or DAK) (Lewis, 2001). More than 85 percent of local road expenditures were financed through DAU, while the remainder came through the DAK. Both are again allocated to local units based on national formulas, where both the criteria and weights changed every few years. The criteria and weights for the DAK are in Figure 7 and for the DAU in Figure 8.[8] Most of the changes in the weights were driven by the decentralization laws and in order to speed up compliance with national priorities of improving equal economic development (Bank, 2007). An additional source of variation in road budgets came from changes in the criteria included in the formula—for instance, the human development index and GDP criteria replaced the poverty index criteria in 2006 (World Bank, 2012). Again, the post-decentralization budgeting process for infrastructure supports the idea of two-stage budgeting.

Direct Central Spending. Although most central government infrastructure expenditures disbursed through the national treasury were probably used for legitimate central government functions, central funds have sometimes been used for subnational responsibilities. Lewis and Chakeri (2004) discuss how national funds may be disbursed directly to a road contractor in the capital, who might use this

[8] For more details on revenue sources over time see Appendix Figure A.5.

to directly upgrade a provincial road. However, such spending was declared to be illegial according to Law 25/1999 on the Financial Balance between the Center and Regional Governments. In robustness checks, we examine whether the two-stage budgeting strategy is less predictive of road improvements before decentralization.

Pre-determined Allocation Formulas. Importantly, both before and after decentralization, the allocation formulas for central to local grants were publicly available. INPRES grant allocation rules were based largely on interregional and intersectoral allocation formulas devised by *Bappenas*, in consultation with the Ministry of Home Affairs and the Ministry of Finance (Crane, 1995). The formulas were set in Jakarta with no local inputs, nor are there annual negotiations with local governments over the allocations. More recently, formulas for DAU were designed to help to equalize fiscal capacities of sub-national governments, to subsidize poorer more remote areas, and are aligned with national priorities stated in *Repelita* documents (Bird and Smart, 2002).

Reliance of Local Governments on Central Transfers. Local governments in Indonesia tend to be heavily reliant on the central government for revenue used for local infrastructure spending, and this fact has not changed much as a consequene of decentralization. Before decentralization, local tax rates were equalized everywhere, and local governments had limited autonomy in their revenue policies (Hill, 1998). The central government maintained control of all major tax bases, even after decentralization, and the bulk of local government revenues came from central government transfers. For instance, Fane (2003) uses Ministry of Finance data to show that in FY 2002, local governments generated just under 5 percent of their total revenue from own-tax and non-tax sources, on average, with the central government accounting for the remaining 95 percent. By the end of 2007, subnational governments accounted for 38 percent of total public sector expenditure but only about 8 percent of total public revenue (Lewis and Oosterman, 2009a; Lewis, 2010). To the extent that local spending on roads was responsive to local economic activity, it was probably for provinces that were large oil and gas producers, and we explore this in robustness checks.

Lack of Borrowing to Finance Road Maintenance. Even before decentralization, regional infrastructure spending was conducted through year-to-year budget allocations, with little borrowing to finance investments and few projects that spanned multiple years (Crane, 1995). Subnationals tend to finance capital spending entirely out of operating balances (cash and reserves), instead of through borrowing (Lewis and Oosterman, 2009b). As a consequence, most infrastructure spending in Indonesia is funded through single-year contracts (93% of contracts signed in 2015 by the Ministry of Public Works and Housing) (Ray and Ing, 2016). After procurement and moblization, there is often little time left for implementation. Construction is rushed, often to substandard quality. The rush to complete projects before the end of the year commonly coincides with the onset of the rainy season, so roads and other assets are often built or repaired in the wet, further undermining quality.

Fixed Road Administration Status. It is also important to note that over the period we study, road administration status remains fairly constant. Appendix Figure A.6 uses BPS data to plot the total length of roads administered by national, provincial, and district governments. Panel A shows total road length for all roads in Indonesia, while Panel B reports the shares of national and provincial roads in total road length. While road length is increasing nationally, provincial roads accounted for 64 percent of total national and provincial road length, on average, from 1990-2003. This figure decreases somewhat in 2004, but IRMS data suggest that those changes occurred outside of Java, Sumatra, and Sulawesi. Panels C and D show that in our sample, provincial roads accounted for 52 percent of total national and provincial road length, a figure that did not fluctuate over time.

Given the weight of the evidence discussed above, temporal variation in local road budgets is plausibly exogenous to changes in local economic activity. While cross-sectional budget allocations may be driven in part by variation in local economic conditions, changes in local budgets over time are not. The variation in changes in local budgets over time is driven by changes in national tax revenues, changes in the criteria and weights in the allocation formulas, and by the fact that different districts have different shares of national and provincial roads and that each of these units has different national budget allocations.

4.2 How Multistage Budgeting Generates Instruments

In this subsection, we explain how this multistage budgeting process can be used to generate instruments for road quality. Let $i = 1, ..., N$ index all districts belonging to a single province, which is itself indexed by p. Each year, the national government allocates a total budget for roads to province p, given by $\overline{B}_{pt} \equiv \overline{B}_t$. After that provincial budget is allocated, provincial officials make road investments by following national formulas for road spending, as follows:

$$B_{it} = \alpha_{it}\overline{B}_t \tag{5}$$

where α_{it} is budget share given by the following:

$$\alpha_{it} = \alpha_{0t} + \mathbf{x}'_{it}\theta_t + \varepsilon_{it} \tag{6}$$

Here, α_{0t} represents a year-specific intercept, and $\mathbf{x}_{it}$ denotes a $(K \times 1)$ vector of observable characteristics of district i at time t. These observable characteristics could include fixed factors, such as fixed geographic characteristics, the area of the district, or historical population measures, but they may also include time-varying factors, including road quality in lagged periods, natural resource revenues in year t, or other tax revenues in year t. The term ε_{it} represents an unobservable factor that shifts the budget share for district i at time t. Importantly, these budget shares will vary over time, owing both to changes in the allocation formulae (as described above) and to changes in district characteristics or unobservables.

Each year, because these budget shares sum to 1, we will have the following:

$$\alpha_{it} = 1 - \sum_{j \neq i} \alpha_{jt}$$
$$\Longrightarrow \alpha_{it} = 1 - (N-1)\,\alpha_{0t} - \sum_{j \neq i} \mathbf{x}'_{jt}\theta_t - \sum_{j \neq i} \varepsilon_{jt}$$

Define $\mathbf{s}_{it} = \left[\sum_{j \neq i} x_{jt}^{(1)}, \sum_{j \neq i} x_{jt}^{(2)}, ..., \sum_{j \neq i} x_{jt}^{(K)}\right]'$ to be a $(K \times 1)$ vector of sums of other districts' observable characteristics that influence budget shares. Using this definition, we can write:

$$\alpha_{jt} = 1 - (N-1)\,\alpha_{0t} - \mathbf{s}_{it}'\theta_t - \sum_{j\neq i}\varepsilon_{jt} \tag{7}$$

So that district i's budget for roads at time t is given by:

$$B_{it} = \left(1 - (N-1)\,\alpha_{0t} - \mathbf{s}_{it}'\theta_t - \sum_{j\neq i}\varepsilon_{jt}\right)\overline{B}_t \tag{8}$$

Equation (8) suggests that district i's road budget at time t depends on the total budget for road improvements in the province, $\overline{B}_t$, interacted with the characteristics of the sum of all other districts in the province, $\mathbf{s}_{it}$. These other district characteristics interacted with $\overline{B}_t$ directly affect district i's budget for road improvements at time t, but because they do not directly affect any outcomes of interest for district i, they should satisfy the exclusion restriction.

In estimation, we also try specifications using a district's own-characteristics ($\mathbf{x}_{jt}$) interacted with $\overline{B}_t$ as IVs. This approach relies more on the exogeneity of $\overline{B}_t$, but if crucial variables for $\mathbf{x}_{jt}$ are not observed, this own-characteristics approach may perform better than the other-characteristics IVs.

4.3 Implementation Details

To make use of (8) to develop instruments, we need to use data to specify two sets of features: (1) a province's total budget for roads at time t, $\overline{B}_t$, and (2) the sum of all other districts' characteristics in the province. We describe each of these features in detail.

Measuring Road Budgets. To measure $\overline{B}_t$, we approximate the total budget for national and provincial roads in the province using the road roughness data.[9] Let $A \in \{N, P\}$ index road maintenance authorities (e.g. National, N, or Provincial, P), and let t index years. Let r index road segments, and let $\mathcal{R}_p^A$ denote the set of road segments in province p under maintenance authority A. We assume that road r was upgraded between $t-1$ and t if its roughness improves, i.e.:

$$U_{rt} = \mathbf{1}\{\mathrm{IRI}_{r,t} < \mathrm{IRI}_{r,t-1}\} \tag{9}$$

Using these upgrading indicators, we measure the total roads upgraded in province p under different maintenance authorities as follows:

$$\widetilde{B}_{p,t}^A = \sum_{r\in\mathcal{R}_p^A} d_r U_{rt} \times (\mathrm{IRI}_{r,t-1} - \mathrm{IRI}_{r,t})$$

where again d_r denotes the length of segment r. In words, the budget for road improvements $\widetilde{B}_t^A$ for administrative authority A in year t equals the total kilometers of roads upgraded in year t that are administered by authority A, weighted by the change in roughness for each road segment. This implies that larger improvements in road quality represent a larger budget spent improving the road. If the costs

[9] Note that we do not use direct data on total amounts allocated to roads, as these may not be so informative about actual road improvements given corruption and monitoring issues common to developing countries (Olken, 2007).

of upgrading are approximately linear, then weighting the upgrading indicator by the change in road roughness should be a good proxy for the total amount of expenditures on roads by the authority. In addition to using $\widetilde{B}^{A}_{p,t}$, we also try proxies that normalize this measure by the total kilometers of roads under that authority in the province (e.g. $\widetilde{B}^{A*}_{p,t} = \widetilde{B}^{A*}_{p,t} / \sum_{r \in \mathcal{R}^A_p} d_r$).

Validity of Road Budgets. For road budgets to be relevant instruments, local budgets should be strong predictors of road quality. To assess this we estimate a version of equation (4). Specifically, we regress a district's road quality against last year's road quality, years since last upgraded in quadratic form, and the budget variables. Table 1 shows that budgets for all types of roads in the province (district) are indeed strongly correlated to road quality in the province (district). The coefficients are of expected sign and show that larger budgets of road authorities imply improved road quality. This confirms the relevance of our instrument for road quality.

For road budgets to satisfy the exclusion restriction, they have to be uncorrelated with local economic conditions. We provide empirical evidence that this is the case by regressing budgets against lagged economic activity. Since budgets for period t are formed in period $t-1$, we regress the budgets against indicators of economic activity in period $t-2$ while including district and year fixed effects. In particular, we regress the road budget of provincial road authorities against the lagged province GDP or lagged number of firms. Table 2 shows that budgets allocated to local authorities are indeed orthogonal to local economic conditions. This increases our confidence that the exclusion restriction of our proposed instrument at the district level is satisfied and the financing of the first stage is indeed compliant with institutional setting described above.

Measuring Own and Other-District Characteristics. To measure $\mathbf{x}_{it}$ and $\mathbf{s}_{it}$, we begin by using several fixed characteristics of districts, many of which are mentioned in the allocation formulae described above. These characteristics, summarized in Appendix Table A.2, can be grouped into four categories: (1) physical characteristics (e.g. area, elevation, slope, ruggedness, distance to major cities); (2) land cover (cultivated, forest, grass, water coverage, built-up land); (3) historical population data; and (4) road characteristics (length of different types of roads, changes in elevation along different types of roads, and slope of different types of roads). We also work with time-varying measures, including district-level measures of total fiscal transfers (DAK, DAU, and DBH SDA), as well as the share of households with access to safe sanitation, safe water, and electricity. These characteristics, along with our measures of $\widetilde{B}^{A}_{p,t}$ and $\widetilde{B}^{A*}_{p,t}$, are summarized in Appendix Table A.1.

Because we work with multiple proxies for $\overline{B}_t$ and multiple measures of $\mathbf{s}_{it}$, we have a potentially large set of instruments from which to choose, and many of these instruments may be weak on their own. As a result, we use post-double-selection lasso techniques to select the appropriate instruments in the analysis, following Belloni et al. (2012).

5 Results

This section presents reduced-form estimates of the impact of road quality on local economic development outcomes. We begin by showing that local road quality improvements reduced travel times between IFLS communities and the nearest provincial or district capitals. Next, we examine the impacts of local road quality improvements on large manufacturing firms. We then study the impact of local road quality improvements on individual and household-level consumption and employment outcomes. Finally, we examine how road quality improvements affected migration and prices.

Road Quality and Travel Times. In the community module of the IFLS, community informants (e.g. village heads in early waves, but in later waves, local leaders, such as school principals, health professionals, religious leaders, or local community organizers) were asked several questions about the travel times between that village and the nearest district or provincial capital. We study the impact of local road improvements on travel times to the nearest provincial or district capital, using the following regression specification:

$$\log y_{ct} = \alpha_c + \alpha_t + \beta \log \text{Road Quality}_{ct} + \mathbf{x}'_{ct}\theta + \varepsilon_{ct} \tag{10}$$

where α_c denotes a community fixed effect, α_t is a survey wave (year) fixed effect, Road Quality_{ct} is our road quality measure for community c's district at time t, $\mathbf{x}_{ct}$ is a vector of controls, and ε_{ct} is an error term. Equation (10) is a log-log specification, so the parameter of interest, β, can be interpreted as an elasticity. Robust standard errors, clustered at the community level, are reported in parentheses.

Because our analysis studies the effects of local road quality improvements on a large set of outcomes, we need to account for multiple hypothesis testing. In all tables, in addition to reporting conventional clustered standard errors, we also report two-stage false-discovery rate (FDR) sharpened q-values in brackets (Benjamini et al., 2006; Anderson, 2008). These q-values represent adjustments to p-values that account for the multiple hypothesis tests we run.

In Table 3, Panel A, we report the results of estimating 10 with log stated travel times (from IFLS) as the dependent variable. In column 1, we report fixed-effects least squares (FELS) estimates, while in column 2, we report the IV-Lasso results for the other-district IVs. The corresponding Kleibergen-Paap Wald Rank F-Stat, a generalization of the first-stage F-statistic for multiple instrumental variables, is reported next to column 2. Column 3 reports the own-district IVs estimates, with the corresponding Kleibergen-Paap Wald Rank F-Stat reported in the next column. Finally, the means of the dependent variables and sample sizes are reported in the last two columns.[10]

Although the other-district IV estimates are not significant in Column 2, we do find a significant negative relationship between road quality and stated travel times to district and provincial capitals in Column 3, using the own-district IV specification. From column 3, a 10 percent increase in road quality causes a 7.3 percent reduction in travel times to the nearest district capital, and a 4.4 percent reduction in travel times to the nearest provincial capital. Both estimates from Column 3 are significant at the 5 percent level on their own and according to the FDR sharpened q-values.

If travel times in the IFLS are noisy and subject to recall bias, estimates from Panel A may be attenuated with measurement error. In Panel B, instead of using stated travel times, we use travel times derived from the underlying roughness data as the dependent variable. To construct this variable, in each year, we selected the nearest provincial or district capital to each IFLS community (based on crow-flies distance) and calculated travel times based on the continuous roughness data.[11] These specifications show larger and more precisely estimated effects of improvements of local road quality on travel times.

[10]The IVs that were selected in both the other-district and own-district IV-lasso specifications are reported in Appendix Table A.3.

[11]See Appendix B.1.4 for more details on the mapping between road roughness and travel speeds.

Another reason for measurement error is that the locations of provincial and district capitals have been changing over time. Until the late 1990s, district boundaries were relatively stable (Booth, 2011).[12] However, as a consequence of the decentralization process, many groups of subdistricts split off from their original districts, forming new districts. The number of districts increased from 302 in 1999 to 514 in 2014, through a process known as *pemekaran* or blossoming (Bazzi and Gudgeon, 2021). This may have created new district capitals in our sample and altered travel times. However, in general, provincial boundaries did not change over this period, nor did the locations of capital cities. In Appendix Table A.4, we drop villages where provincial capitals changed, and we find similar results, although now the own-district IV-lasso specifications are significant.[13] In summary, using our IV strategy, we find that road quality improvements lead to significant reductions in travel times between IFLS communities and the nearest large cities, as proxied by district and provincial capitals.

Road Quality and Manufacturing Outcomes. We next turn to an investigation of the impact of road quality on manufacturing outcomes. To do so, we began by creating district-year aggregates of the individual firm-level data from the SI. We then used these district-year variables in a panel regression to understand the relationship between road quality and manufacturing outcomes. The regression equation we used is the following:

$$y_{dt} = \alpha_d + \alpha_t + \beta \log \text{Road Quality}_{dt} + \mathbf{x}'_{dt}\theta + \varepsilon_{dt} \tag{11}$$

where α_d and α_t are district and year fixed effects respectively, $\log \text{Road Quality}_{dt}$ is district d's log average road quality measure in year t, $\mathbf{x}_{dt}$ are time-varying controls, and ε_{dt} is the error term. Robust standard errors, clustered at the district level, are reported in parentheses.

Table 4 shows the results, with a structure similar to Table 3. Overall, the Kleibergen-Paap Wald Rank F statistics for both sets of IVs are large, suggesting that our IV models are well specified.[14] In Panel A, we describe the effects of road quality on the existence and quantity of manufacturing firms in the district. The first row uses an indicator for whether or not there are any manufacturing firms in the district as the dependent variable. Although the fixed-effects least squares estimates in Panel A tend to be small and insignificant, we report a positive and significant effect of road quality on the presence of manufacturing firms in the district in both IV-Lasso specifications (columns 2 and 3). This suggests that naive estimates of treatment effects of the impact of road quality may suffer either from negative targeting bias (e.g. policymakers upgraded roads in less developed areas) or from negative feedback (e.g. faster growing areas had greater road deterioration).

In row 1, the IV-lasso estimates show that a 10 percent increase in road quality leads to a 0.7-0.9 percent increase in the probability of firms being in the district. This effect is significant, but economically small, unsurprisingly given that over 95 percent of district-years in the sample have at least 1 firm. From the IV-Lasso specifications in row 2, we find that a 10 percent increase in road quality is associated with a 6 percent increase in the number of new firm openings in the district. Taken literally, given the

[12]Many district boundaries originated under colonial rule when the Dutch relied on local leaders to implement indirect rule.

[13]In Appendix Table A.5, we use island-market potential instead of road roughness as the key independent variable, and obtain similar results.

[14]The IVs that were selected in both the other-district and own-district IV-lasso specifications are reported in Appendix Table A.10.

mean number of new firms across district-years, a new firm would emerge in the district if road quality improved by 30 percent.

In row 3, we also find mixed evidence that road quality improvements lead to a reduction in firm closures, although estimates differ somewhat between the different IV specifications. Finally, in row 4, we show that a 10 percent increase in road quality leads to a 1.2-1.6 percent change in the number of firms in the district. In summary, the results from Table 4, Panel A show that road quality improvements lead to moderately sized, positive effects of the number of firms in the district, and these effects are driven mostly by new firms rather than firm closures or expansion into new areas.

In Panel B, we show the effect of road quality improvements on district-level production outcomes. In row 5, the IV-lasso specifications show that a 10 percent increase in road quality leads to a 5-8 percent in district-level value added. In row 6, we find that a 10 percent increase in road quality leads to a 4-7 percent increase in value added. However, we find somewhat smaller effects of road quality improvements on the total number of manufacturing workers in large firms in the district. A 10 percent increase in road quality leads to a 2.5-2.7 percent increase in the number of manufacturing workers in the district.

Because the effects of road quality on output are larger than the effects on employment, it is not surprising that that we see positive effects on output per worker. A 10 percent increase in road quality leads to a 3-3.5 percent increase in output per worker for large manufacturing firms in the district.[15]

To what extent are the manufacturing results from Table 4 explained by intensive-margin improvements of existing firms? An advantage of the SI data is that they contain firm-level identifiers, so we can study how existing firms production outcomes changed in responses to changes in road improvements. In Table 5, we use our IV strategy to estimate the following firm-level panel regression specification:

$$y_{idt} = \alpha_i + \alpha_t + \beta \log \text{Road Quality}_{dt} + \mathbf{x}'_{dt}\theta + \varepsilon_{dt} \tag{12}$$

where i indexes firms, α_i is a firm fixed effect, and α_t is a year fixed effect. Because we do not observe firms that move in our sample (and in fact, BPS creates new firm identifiers when firms move), the firm fixed effect also controls for any time-invariant district-specific characteristics. We estimate (15) using more than 250,000 firm $\times$ year observations over the 1990-2007 period, and standard errors are again clustered at the district level.

Table 5 shows results on firm-level output (row 1), value added (row 2), total employment (row 3), and output per worker (row 4), mirroring the production outcomes shown in Table 4.[16] Despite large Kleibergen-Paap Wald Rank F statistics for both sets of IVs, we find that there were no significant effects of the road quality improvements on any of these production outcomes. Taken together with evidence from Table 4, our results suggest that road quality improvements affect the extensive margin of firm creation, but they do not have significant effects on pre-existing firms.[17]

Road Quality and Consumption, Income, and Employment. Given that we have observed that road quality improvements lead to an increase in the number of manufacturing jobs, we now turn to

[15] In Appendix Table A.11, we use island-market potential instead of road roughness as the key independent variable, and obtain similar results.

[16] The IVs that were selected in both the other-district and own-district IV-lasso specifications are reported in Appendix Table A.18.

[17] In Appendix Table A.19 we find similar null results of the impact of island-market potential on firm-level production outcomes.

understanding how these new employment opportunities impacted consumption and income. To do so, we use the consumption, income, and labor market data from the IFLS, which tracks roughly 9,500 households and 44,000 individuals over 4 survey waves.

We estimate household (or individual) fixed effects regressions of the following form:

$$y_{idt} = \alpha_i + \alpha_t + \beta \log \text{Road Quality}_{dt} + \mathbf{x}'_{idt}\theta + \varepsilon_{dt} \tag{13}$$

where α_i denotes a household (or individual) fixed effect and α_t denotes a survey wave fixed effect, as before. Included in $\mathbf{x}_{idt}$ are controls for current levels of population, GDRP, survey month indicators, and controls for household size. When we run individual-level regressions, we also add time-varying controls for individual age and education.

In Panel A, which uses household-level fixed effects regressions, we estimate positive and statistically significant effects of increases in road quality on consumption.[18] We find that a 10 percent increase in road quality increases household consumption expenditures by 1.8-2 percent. In Panel B, we use individual-level fixed effects regressions to show that a 10 percent increase in road quality leads to a 2 percent increase in total earnings. We also find that road quality improvements increased agricultural income; a 10 percent increase in road quality increased agricultural incomes by 2.5-4.4 percent.

In the next set of rows for Panel B, we examine the impact of road quality improvements on hours worked and employment outcomes. We first show that road quality improvements do not have any impact on the total number of hours worked (row 4) or on the probability of being employed (row 5). However, in the final rows of Table 6, we show that the effects on log total earnings and consumption come from sector switching.

We first assigned every employed worker into one of three different employment sectors: (1) manufacturing; (2) other formal employment; and (3) other informal employment. We find that a 10 percent increase in road quality leads to a 0.7-1.8 percent increase in the probability of working in manufacturing (row 6). We also find that there was no effect on the probability of working in the formal sector (row 7). However, across both sets of IV specifications, we find that road quality improvements led workers to exit the informal sector (row 8). A 10 percent increase on road quality leads to a 1.5-1.8 percent reduction in the probability of working in the formal sector, conditional on working.

These results represent key findings of the paper and represents the household employment counterpart to the findings on the manufacturing sector in Tables 4 and 5. The manufacturing employment effects that we see from Table 4 seem to be coming from workers moving out of the informal sector (some of which includes agricultural employment), consistent with road quality playing a crucial role in local economic development and structural transformation.[19]

Road Quality and District-Level Migration Outcomes. Given that road quality improvements were associated with increased in-migration of firms and better employment opportunities outside of the informal sector, we would expect to see some migratory responses. To study how local road quality improvements impacted migration, we use census data from 1990 and 2000 to examine how the number of

[18] The IVs that were selected in both the other-district and own-district IV-lasso specifications are reported in Appendix Table A.24.

[19] In Appendix Table A.25 we find similar null results of the impact of island-market potential on firm-level production outcomes.

recent migrants (i.e. those arriving within the last 5 years) was impacted by changes in road roughness.

Our empirical specification is the following:

$$y_d = \alpha_{p(d)} + \beta \Delta \log \text{Road Quality}_d + \mathbf{x}_d' \theta + \varepsilon_d \ , \tag{14}$$

where y_d is a cross-sectional migration outcome, $\alpha_{p(d)}$ represents a fixed effect for district d's province, $\Delta \log \text{Road Quality}_d$ is the change in log road quality between 1995 and 2000, and $\mathbf{x}_d$ is a vector of additional controls, which often includes logs of population and GDP in 1990. In this specification, instead of instrumenting road quality in levels, we instrument the change in road quality between 2000 and 1995 with the same IV-lasso strategy as above.

The first row of Table 7 shows that districts with improved road quality experienced a significantly more rapid rate of population growth. The first row reports an elasticity, so that a 10 percent increase in road quality from 1995 to 2000 is associated with a 3.9 percent increase in 10-year population growth.[20]

In the next two rows, we regress the log total recent migrants from different districts and different provinces on the change in road quality. We find that a 10 percent increase in road quality growth leads to a 7.9 percent increase in the number of district migrants and a 4.1 percent increase in the number of province migrants, although the latter effect is not significant. In numerical terms, the effect on district migrants is modest, with a 10 percent increase in road quality leading to roughly an additional 3,400 migrants for the average district. The average district has a population of 726,979 in 2000, so this increase would represent less than one half of 1 percent of the population.

Finally, in row 4, we use province-migration data from 1990 to show that a 10 percent increase in road quality leads to a 3.3 percent increase in the 10-year growth rate of province migrants. Although these cross-sectional estimates are less well identified than the panel specifications used in the rest of the paper, they nevertheless provide evidence that road quality improvements modestly increase migration.

Road Quality and Prices. In this paper, we argue that local road quality can represent an amenity valued by both producers and consumers. From our discussion of the model presented above, a simple spatial equilibrium model would suggest that if migration costs are small, firms and workers would move to upgraded locations in response to increases in road quality. Greater in-migration should increase land and housing prices, but the impacts on wages depend on the relative shifts of the demand and supply of labor (Rosen, 1979; Roback, 1982). We have already documented a modest migratory response of both firms and workers, so we now investigate the impact on housing and other prices.

Our regression specification is as follows:

$$y_{cdt} = \alpha_c + \alpha_t + \beta \log \text{Road Quality}_{dt} + \mathbf{x}_{cdt}' \theta + \varepsilon_{cdt} \tag{15}$$

where c indexes IFLS communities, α_c is a community fixed effect, and α_t is a year fixed effect.

Table 8 shows our results. In rows 1 and 2, we show that local road quality improvements are not significantly correlated with factory or farm wages. Although the coefficients on factory wages are large, they are not significant, whereas the coefficients on farm wages seem to be more precisely estimated zeros.

[20] The IVs that were selected in both the other-district and own-district IV-lasso specifications are reported in Appendix Table A.31.

In rows 3 and 4, we study the effect of local road quality on food prices. The food price measure we use is a Laspeyres price index composed of community prices of tradeable goods (including rice, oil, sugar, and salt) and perishable goods (including meat and fish). Corresponding initial consumption values are used for expenditure weights. In row 3, we show no clear correlations between road quality and overall food prices, but we do see a significant negative effect on perishable food prices in row 4. A 10 percent increase in road quality leads to a 3.7-6.4 percent reduction in perishable food prices.

Finally, in the last four rows of the table, we study the relationship between road quality and housing prices. We use both log median land values (row 5) and log median rents (row 6) as well as estimates of log land values and log rents from hedonic specifications. For the hedonic specifications, we first estimate a hedonic price regression of log rents or log land values on a large vector of household and plot characteristics, in addition to controlling for fixed effects at the community-by-wave level.[21] We use these estimated community-by-wave fixed effects as the dependent variable in rows 7 and 8. Overall, although the log rent specifications are noisy and not significant, we tend to find a positive relationship between road quality improvements and land values. However, these effects are relatively imprecisely estimated.

6 Conclusion

Even though road maintenance investments typically account for a significant proportion of countries' budgets, little is known about their effects in developing countries, where spatial disparities are particularly pronounced. This paper aims to understand the role that road improvement (or deterioration) can play in such countries, not only through looking at possible welfare effects, but also by investigating the different possible mechanisms through which these effects materialize. While much of the previous literature on this topic has focused on the construction of new roads, we add to the literature by evaluating the effects of substantial changes in road quality due to maintenance and upgrading of existing national and provincial highways in Indonesia.

Using a novel dataset that documents substantial variation in road quality in Indonesia, and combining this with high quality household panel data from 1990 to 2007, we provide reduced form evidence that road improvements significantly increase welfare, measured either with consumption or income. Using an annual census of manufacturing firms, we show that these positive welfare effects partly materialize though increased labor market demand, generated by the entry of new firms rather than through hiring by existing firms. We do not see substantial changes in the extensive or intensive margin of labor supply, but instead observe occupational shifts from informal employment (including agriculture) into higher paying, newly available manufacturing jobs. In addition, while manufacturing wages typically don't exhibit an upward push, we do observe significant improvements in agricultural profits. This not only implies the wage gap between these two sectors is narrowed, but also confirms the predictions of our stylized model of dual labor markets. The latter shows under what conditions productive amenities, such as transport infrastructure, may translate into positive welfare effects.

The methodological contribution of this paper is in addressing the common concerns of targeting bias and reverse causality by suggesting a new instrument, replicable in many instances. We take advan-

[21] Estimates of the hedonic relationships can be found in Appendix Table A.32.

tage of Indonesia's institutional two-step budgeting setup for road funding, where different authorities, such as provinces or districts, are in charge of road quality and funding of different parts of the road network. This allows us to construct a time varying instrument for road quality based on allocation formulas. Thus, we identify the effects from the set of roads that get maintained when road budgets allow for it, but which get less maintenance when road budgets are tight or scaled back.

The evidence presented in this paper shows that improving major national and provincial roads can modestly improve local economic development through increasing formal labor market opportunities. Conversely, deterioration of these important roads may have adverse effects in the opposite direction. Governments should be aware of the impacts of road maintenance investments when setting priorities for transportation budgets.

References

AGGARWAL, S. (2018): "Do Rural Roads Create Pathways out of Poverty? Evidence from India," *Journal of Development Economics*, 133, 375–395.

AHMAD, E. AND A. M. MANSOOR (2002): *Indonesia: managing decentralization*, International Monetary Fund.

ANDERSON, M. L. (2008): "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American statistical Association*, 103, 1481–1495.

ASHER, S. AND P. NOVOSAD (2020): "Rural roads and local economic development," *American Economic Review*, 110, 797–823.

ASIA FOUNDATION (2008): "The Cost of Moving Goods: Road Transportation, Regulations and Charges in Indonesia," Survey report.

ASTURIAS, J., M. GARCÍA-SANTANA, AND R. RAMOS (2019): "Competition and the welfare gains from transportation infrastructure: Evidence from the Golden Quadrilateral of India," *Journal of the European Economic Association*, 17, 1881–1940.

BANERJEE, A., E. DUFLO, AND N. QIAN (2012): "On the Road: Access to Transportation Infrastructure and Economic Growth in China," Working Paper.

BANK, W. (2007): "Spending for Development: Making the Most of Indonesia's New Opportunities," Tech. rep.

BARDHAN, P. AND C. UDRY (1999): *Development Microeconomics*, Oxford: Oxford University Press.

BAUM-SNOW, N. (2007): "Did Highways Cause Suburbanization?" *Quarterly Journal of Economics*, 122, 775–805.

BAZZI, S. AND M. GUDGEON (2021): "The political boundaries of ethnic divisions," *American Economic Journal: Applied Economics*, 13, 235–66.

BELLONI, A., D. CHEN, V. CHERNOZHUKOV, AND C. HANSEN (2012): "Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain," *Econometrica*, 80, 2369–2429.

BENJAMIN, D. (1992): "Household Composition, Labor markets, and Labor demand: Testing for Separation in Agricultural Household Models," *Econometrica*, 60, 287–322.

BENJAMINI, Y., A. M. KRIEGER, AND D. YEKUTIELI (2006): "Adaptive linear step-up procedures that control the false discovery rate," *Biometrika*, 93, 491–507.

BENNETT, C. R., A. CHAMORRO, C. CHEN, H. DE SOLMINIHAC, AND G. W. FLINTSCH (2007): "Data Collection Technologies for Road Management," Technical report, East Asia Pacific Transport Unit, World Bank.

BIRD, J. AND S. STRAUB (2020): "The Brasilia experiment: the heterogeneous impact of road access on spatial development in Brazil," *World Development*, 127, 104739.

BIRD, R. M. AND M. SMART (2001): "Intergovernmental fiscal transfers: Some lessons from international experience," in *March (revision). Paper prepared for Symposium on Intergovernmental Transfers in Asian Countries: Issues and Practices, Asian Tax and Public Policy Program, Hitotsubashi University*.

——— (2002): "Intergovernmental fiscal transfers: International lessons for developing countries," *World Development*, 30, 899–912.

BLALOCK, G. AND P. J. GERTLER (2008): "Welfare Gains from Foreign Direct Investment Through Technology Transfer to Local Suppliers," *Journal of International Economics*, 74, 402–421.

BLIMPO, M. P., R. HARDING, AND L. WANTCHEKON (2013): "Public investment in rural infrastructure: Some political economy considerations," *Journal of African Economies*, 22, ii57–ii83.

BOCK, M., A. CARDAZZI, AND B. R. HUMPHREYS (2021): "Where the Rubber Meets the Road: Pavement Damage Reduces Traffic Safety and Speed," NBER Working Paper 29176.

BOOTH, A. (2003): "Decentralisation and Poverty Alleviation in Indonesia," *Environment and Planning C: Government and Policy*, 21, 181–202.

——— (2011): "Splitting, splitting and splitting again: A brief history of the development of regional government in Indonesia since independence," *Bijdragen tot de taal-, land-en volkenkunde/Journal of the Humanities and Social Sciences of Southeast Asia*, 167, 31–59.

BRYAN, G. AND M. MORTEN (2019): "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia," *Journal of Political Economy*, 127, 2229–2268.

BURGESS, R., R. JEDWAB, E. MIGUEL, A. MORJARIA, AND G. PADRÓ I MIQUEL (2015): "The Value of Democracy: Evidence from Road Building in Kenya," *American Economic Review*, 105, 1817–51.

CRANE, R. (1995): "The practice of regional development in Indonesia: Resolving central-local coordination issues in planning and finance," *Public Administration and Development*, 15, 139–149.

DEATON, A. AND J. MUELLBAUER (1980): "An almost ideal demand system," *The American economic review*, 312–326.

DICK, H. (2000): "Representations of development in 19th and 20th century Indonesia: a transport history perspective," *Bulletin of Indonesian Economic Studies*, 36, 185–207.

DONALDSON, D. (2018): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, 108, 899–934.

DONALDSON, D. AND R. HORNBECK (2016): "Railroads and American Economic Growth: A 'Market Access Approach," *Quarterly Journal of Economics*, 131, 799–858.

DURANTON, G. AND M. A. TURNER (2012): "Urban Growth and Transportation," *The Review of Economic Studies*, 79, 1407–1440.

FABER, B. (2014): "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System," *Review of Economic Studies*, 81, 1046–1070.

FANE, G. (2003): "Change and Continuity in Indonesia's New Fiscal Decentralisation Arrangements," *Bulletin of Indonesian Economic Studies*, 39, 159–176.

GHANI, E., A. G. GOSWAMI, AND W. R. KERR (2016): "Highway to success: The impact of the Golden Quadrilateral project for the location and performance of Indian manufacturing," *The Economic Journal*, 126, 317–357.

GOLLIN, D. AND R. ROGERSON (2014): "Productivity, transport costs and subsistence agriculture," *Journal of Development Economics*, 107, 38–48.

HILL, H. (1998): "The challenge of regional development in Indonesia," *Australian Journal of International Affairs*, 52, 19–34.

JACOBY, H. G. (2000): "Access to Markets and the Benefits of Rural Roads," *The Economic Journal*, 110, 713–737.

KHANDKER, S. R., Z. BAKHT, AND G. B. KOOLWAL (2009): "The Poverty Impact of Rural Roads: Evidence from Bangladesh," *Economic Development and Cultural Change*, 57, 685–722.

KHANDKER, S. R. AND G. B. KOOLWAL (2011): "Estimating the long-term impacts of rural roads: a dynamic panel approach," World Bank Policy Research Working Paper 5867.

LA PORTA, R. AND A. SHLEIFER (2014): "Informality and Development," *Journal of Economic Perspectives*, 28, 109–26.

LEIGLAND, J. (1993): "Decentralizing the Development Budget Process in Indonesia: Progress and Prospects," *Public Budgeting & Finance*, 13, 85–101.

LEWIS, B. D. (2001): “The new Indonesian equalisation transfer,” *Bulletin of Indonesian Economic Studies*, 37, 325–343.

——— (2010): “Indonesian decentralization: Accountability deferred,” *International Journal of Public Administration*, 33, 648–657.

——— (2017): “Does local government proliferation improve public service delivery? Evidence from Indonesia,” *Journal of Urban Affairs*, 39, 1047–1065.

LEWIS, B. D. AND J. CHAKERI (2004): “Central Development Spending in the Regions Post-Decentralisation,” *Bulletin of Indonesian Economic Studies*, 40, 379–394.

LEWIS, B. D. AND A. OOSTERMAN (2009a): “The Impact of Decentralization on Subnational Government Fiscal Slack in Indonesia,” *Public Budgeting & Finance*, 29, 27–47.

——— (2009b): “The Impact of Decentralization on Subnational Government Fiscal Slack in Indonesia,” *Public Budgeting & Finance*, 29, 27–47.

LEWIS, W. A. (1954): “Economic Development with Unlimited Supplies of Labour,” *The Manchester School*, 22, 139–191.

MICHAELS, G. (2008): “The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System,” *The Review of Economics and Statistics*, 90, 683–701.

MORTEN, M. AND J. OLIVERIA (2018): “The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City,” NBER Working Paper 22158.

OLKEN, B. (2007): “Monitoring Corruption: Evidence from a Field Experiment in Indonesia,” *Journal of Political Economy*, 115, 200–249.

RAY, D. AND L. Y. ING (2016): “Survey of Recent Developments: Addressing Indonesia’s Infrastructure Deficit,” *Bulletin of Indonesian Economic Studies*, 52, 1–25.

REDDING, S. J. AND M. A. TURNER (2015): “Chapter 31 Transportation Costs and the Spatial Organization of Economic Activity,” in *Handbook of Urban and Regional Economics, Volume 5B*, ed. by G. Duranton, J. V. Henderson, and W. C. Strange, Elsevier, 1339–1398.

ROBACK, J. (1982): “Wages, Rents, and the Quality of Life,” *Journal of Political Economy*, 90, 1257–1278.

ROSEN, S. (1979): “Wage-Based Indexes of Urban Quality of Life,” in *Current Issues in Urban Economics*, ed. by P. Mieszkowski and M. Straszheim, Johns Hopkins University Press.

SHAH, A., Z. QURESHI, B. BINDER, AND H.-F. ZOU (1994): “Intergovernmental fiscal relations in Indonesia,” World bank discussion paper.

SINGH, I., L. SQUIRE, AND J. STRAUSS (1986): *Agricultural Household Models: Extensions, Applications, and Policy*, Baltimore: Johns Hopkins University Press.

STOREYGARD, A. (2016): “Farther on Down the Road: Transport Costs, Trade and Urban Growth in sub-Saharan Africa,” *The Review of Economic Studies*, 83, 1263–1295.

TEMPLE, J. (2005): “Dual economy models: A primer for growth economists,” *The Manchester School*, 73, 435–478.

VALDIVIA, M. (2011): “Contracting the road to development: early impacts of a rural roads program,” *Poverty and Economic Policy Network PMMA Working Paper*.

WORLD BANK (2008): *Spending for Development: Making the Most of Indonesia’s New Opportunities*, World Bank.

——— (2012): *Investing in Indonesia’s roads : improving efficiency and closing the financing gap*.

YU, J., E. CHOU, AND J. YAU (2006): “Development of Speed-Related Ride Quality Thresholds using International Roughness Index,” *Transportation Research Record*, 1974, 47–53.

Table 1: ROAD ROUGHNESS AND BUDGETS

Dep. Var.: Road segment log IRI(t)	(1)	(2)	(3)	(4)	(5)	(6)
Budget (National Roads)	-0.044		-0.029	-0.047		-0.040
	(0.001)***		(0.001)***	(0.001)***		(0.001)***
Budget (Provincial Roads)		-0.037	-0.027		-0.026	-0.014
		(0.000)***	(0.001)***		(0.000)***	(0.000)***
$\log \text{IRI}_{t-1}$				0.415	0.411	0.414
				(0.002)***	(0.002)***	(0.002)***
Years Since Last Upgrade				0.164	0.165	0.163
				(0.001)***	(0.001)***	(0.001)***
Years Since Last Upgrade^2				-0.012	-0.012	-0.012
				(0.000)***	(0.000)***	(0.000)***
N	955214	960837	950458	766335	769723	762638
Adjusted R^2	0.078	0.079	0.081	0.388	0.386	0.388
F Statistic	2520.469	2734.166	2473.409	11979.469	12074.867	11528.778
Road FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: GDP and # firms are measured at the district level. Robust standard errors in parentheses, clustered at the road segment level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 2: ROAD BUDGETS AND LOCAL ECONOMIC CONDITIONS

	Province Panel			
	Nat. roads	Nat. roads	Prov. roads	Prov. roads
Dep. Var.: $\log \text{Budget}_t$	**(1)**	**(2)**	**(3)**	**(4)**
$\log \text{GDP}_{t-2}$	0.241		0.251	
	(0.269)		(0.406)	
$\log \#\text{Firms}_{t-2}$		0.011		-0.004
		(0.007)		(0.003)
N	241	241	249	249
Adjusted R^2	0.586	0.589	0.643	0.645
Province FE	Yes	Yes	Yes	Yes
District FE				
Year FE	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses, clustered at the road segment level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 3: Road Quality and Travel Times

	FELS	Other District IV-Lasso		Own District IV-Lasso		Stats	
Panel A: IFLS Travel Times	**(1)**	**(2)**	**KBP**	**(3)**	**KBP**	$\overline{Y}$	N
Log Travel Time to Nearest District Capital	-0.092	-0.346	76.904	-0.733**	22.601	46.241	850.000
	(0.145)	(0.298)	(.)	(0.326)	(.)	(.)	(.)
		[0.17]		[0.04]			
Log Travel Time to Nearest Provincial Capital	-0.203***	-0.221	62.470	-0.440**	33.600	172.334	872.000
	(0.073)	(0.138)	(.)	(0.212)	(.)	(.)	(.)
		[0.12]		[0.04]			
Panel B: Roughness-Based Travel Times							
Log Travel Time to Nearest District Capital	-0.487***	-0.731***	112.423	-0.503***	38.405	21.272	904.000
	(0.049)	(0.107)	(.)	(0.101)	(.)	(.)	(.)
		[0.00]		[0.00]			
Log Travel Time to Nearest Provincial Capital	-0.447***	-0.968***	112.423	-1.016***	38.405	75.217	904.000
	(0.032)	(0.094)	(.)	(0.098)	(.)	(.)	(.)
		[0.00]		[0.00]			
Year FE	Yes	Yes		Yes			
Village FE	Yes	Yes		Yes			

Notes: We report the results of community-level panel regressions of the dependent variable on road quality (both in logs). Each cell reports β from a separate regression, with the dependent variable listed in the row heading. All regressions include community and year fixed effects, with controls including logs of current population and non-oil GDRP. Interpretation of results remains unchanged when dependent variables are expressed in levels. Dependent variable means are reported in levels. Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 4: ROAD QUALITY AND DISTRICT-LEVEL MANUFACTURING OUTCOMES

	FELS	Other District IV-Lasso		Own District IV-Lasso		Stats	
Panel A: Firm Counts	**(1)**	**(2)**	**KBP**	**(3)**	**KBP**	$\overline{Y}$	N
Any Firms (0 1)	0.007	0.073*	69.017	0.086**	87.203	0.956	3381.000
	(0.017)	(0.039)	(.)	(0.034)	(.)	(.)	(.)
		[0.08]		[0.02]			
Log Number of Opened Firms	0.058	0.602***	69.017	0.589***	87.203	6.075	3381.000
	(0.060)	(0.134)	(.)	(0.139)	(.)	(.)	(.)
		[0.00]		[0.00]			
Log Number of Closed Firms	-0.042	-0.089	71.135	-0.331*	91.130	6.113	3184.000
	(0.070)	(0.194)	(.)	(0.184)	(.)	(.)	(.)
		[0.12]		[0.03]			
Percent Δ Number of Firms	-0.003	0.120***	69.360	0.161***	69.705	-0.032	3337.000
	(0.020)	(0.044)	(.)	(0.042)	(.)	(.)	(.)
		[0.05]		[0.00]			
Panel B: Production							
Log Output	0.118	0.526**	69.017	0.805***	87.203	1591.043	3381.000
	(0.144)	(0.244)	(.)	(0.249)	(.)	(.)	(.)
		[0.06]		[0.01]			
Log Value Added	0.078	0.434*	69.017	0.705***	87.203	575.353	3381.000
	(0.140)	(0.244)	(.)	(0.254)	(.)	(.)	(.)
		[0.09]		[0.01]			
Log Total Employment	-0.169**	0.253**	69.017	0.271**	87.203	13732.521	3381.000
	(0.081)	(0.120)	(.)	(0.130)	(.)	(.)	(.)
		[0.06]		[0.02]			
Log Output per Worker	0.296***	0.304**	67.035	0.351***	87.297	0.081	3232.000
	(0.086)	(0.136)	(.)	(0.130)	(.)	(.)	(.)
		[0.06]		[0.01]			
Year FE	Yes	Yes		Yes			
District FE	Yes	Yes		Yes			

Notes: We report the results of district-level panel regressions of the dependent variable on road quality. Each cell reports β from a separate regression, with the dependent variable listed in the row heading. All regressions include district and year fixed effects, with controls that include logs of current population and non-oil GDRP. Robust standard errors, clustered at the district level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 5: ROAD QUALITY AND FIRM-LEVEL MANUFACTURING OUTCOMES

	FELS	Other District IV-Lasso		Own District IV-Lasso		Stats	
	(1)	(2)	KBP	(3)	KBP	$\overline{Y}$	N
Log Output	-0.031	-0.004	34.667	0.006	55.001	19.936	278434.000
	(0.041)	(0.081)	(.)	(0.065)	(.)	(.)	(.)
		[1.00]		[1.00]			
Log Value Added	-0.055	0.045	34.637	0.031	54.992	7.371	278368.000
	(0.047)	(0.089)	(.)	(0.068)	(.)	(.)	(.)
		[1.00]		[1.00]			
Log Total Labor	-0.007	-0.018	34.650	-0.031*	54.980	164.259	278539.000
	(0.012)	(0.022)	(.)	(0.019)	(.)	(.)	(.)
		[1.00]		[0.77]			
Log Output per Worker	-0.023	-0.026	34.665	0.031	54.978	0.073	278284.000
	(0.039)	(0.079)	(.)	(0.062)	(.)	(.)	(.)
		[1.00]		[1.00]			
Year FE	Yes	Yes		Yes			
District FE	Yes	Yes		Yes			
Firm FE	Yes	Yes		Yes			

Notes: We report the results of firm-level panel regressions of the dependent variable on road quality. Each cell reports β from a separate regression, with the dependent variable listed in the row heading. All regressions include firm and year fixed effects (and implicitly also district fixed effects), with controls that include logs of current population and non-oil GDRP. Robust standard errors, clustered at the district level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 6: Effects of Road Quality on Consumption, Income, and Employment

	FELS	Other District IV-Lasso		Own District IV-Lasso		Stats	
Panel A: Household-Level Outcomes	**(1)**	**(2)**	**KBP**	**(3)**	**KBP**	$\overline{Y}$	N
Log HH Per-Capita Consumption Expenditures	0.132***	0.203**	150.232	0.183**	53.727	11.068	22036.000
	(0.042)	(0.087)	(.)	(0.081)	(.)	(.)	(.)
		[0.04]		[0.08]			
Panel B: Individual-Level Outcomes							
Log Total Earnings	0.070**	0.232***	120.181	0.082	34.720	1.438	17619.000
	(0.033)	(0.055)	(.)	(0.060)	(.)	(.)	(.)
		[0.00]		[0.19]			
Agriculture ... Log Earnings	0.160**	0.435***	175.484	0.251*	43.961	0.718	5246.000
	(0.071)	(0.122)	(.)	(0.146)	(.)	(.)	(.)
		[0.01]		[0.15]			
Log Total Hours Worked	-0.004	-0.070	182.406	0.016	47.281	199.077	22931.000
	(0.038)	(0.077)	(.)	(0.088)	(.)	(.)	(.)
		[0.29]		[0.34]			
Any Employment (0 1)?	-0.025	-0.030	101.357	0.067	29.351	0.701	36257.000
	(0.021)	(0.037)	(.)	(0.045)	(.)	(.)	(.)
		[0.29]		[0.18]			
Manufacturing ... Any Employment (0 1)?	0.084***	0.071	182.545	0.175***	47.344	0.290	22934.000
	(0.024)	(0.047)	(.)	(0.053)	(.)	(.)	(.)
		[0.13]		[0.01]			
Other, Formal ... Any Employment (0 1)?	0.014	0.040	182.545	0.001	47.344	0.314	22934.000
	(0.021)	(0.052)	(.)	(0.056)	(.)	(.)	(.)
		[0.29]		[0.34]			
Other, Informal ... Any Employment (0 1)?	-0.095***	-0.154***	182.545	-0.181***	47.344	0.465	22934.000
	(0.022)	(0.046)	(.)	(0.050)	(.)	(.)	(.)
		[0.01]		[0.01]			
Individual FE	Yes	Yes		Yes			
Year FE	Yes	Yes		Yes			

Notes: We report the results of individual-level panel regressions with individual and survey-wave fixed effects. Each cell reports estimates of β from a separate regression, with the dependent variable listed in the row heading. Controls include: district GDP, individual age, education, household size, and month of survey indicators. Total hours worked is defined only if the individual reported working. Earnings regressions also include hours worked (by sector) as a control. Robust standard errors in parentheses, clustered at the (initial) village level. */**/*** denotes significant at the 10% / 5% / 1% levels. Full results, including results restricted to the sample of non-moving individuals, can be found in Appendix Tables.

Table 7: Road Quality and District-Level Migration Outcomes

	FELS	Other District IV-Lasso		Own District IV-Lasso		Stats	
	(1)	(2)	KBP	(3)	KBP	$\overline{Y}$	N
Percent Δ Population (2000-1990)	0.053 (0.039)	0.393* (0.232) [0.08]	35.196 (.)	0.201** (0.098) [0.12]	44.872 (.)	0.137 (.)	198.000 (.)
Log Total Recent Migrants (Kabu)	0.427** (0.167)	0.792** (0.332) [0.06]	32.503 (.)	1.213** (0.593) [0.12]	29.457 (.)	43514.102 (.)	198.000 (.)
Log Total Recent Migrants (Prov)	0.208 (0.281)	0.414 (0.392) [0.17]	32.503 (.)	-0.037 (0.601) [0.31]	29.457 (.)	20453.100 (.)	198.000 (.)
Percent Δ Prov. Migrants (2000-1990)	0.095** (0.043)	0.326** (0.127) [0.06]	24.795 (.)	0.197* (0.105) [0.12]	61.454 (.)	-0.828 (.)	181.000 (.)
Province FE	Yes	Yes		Yes			

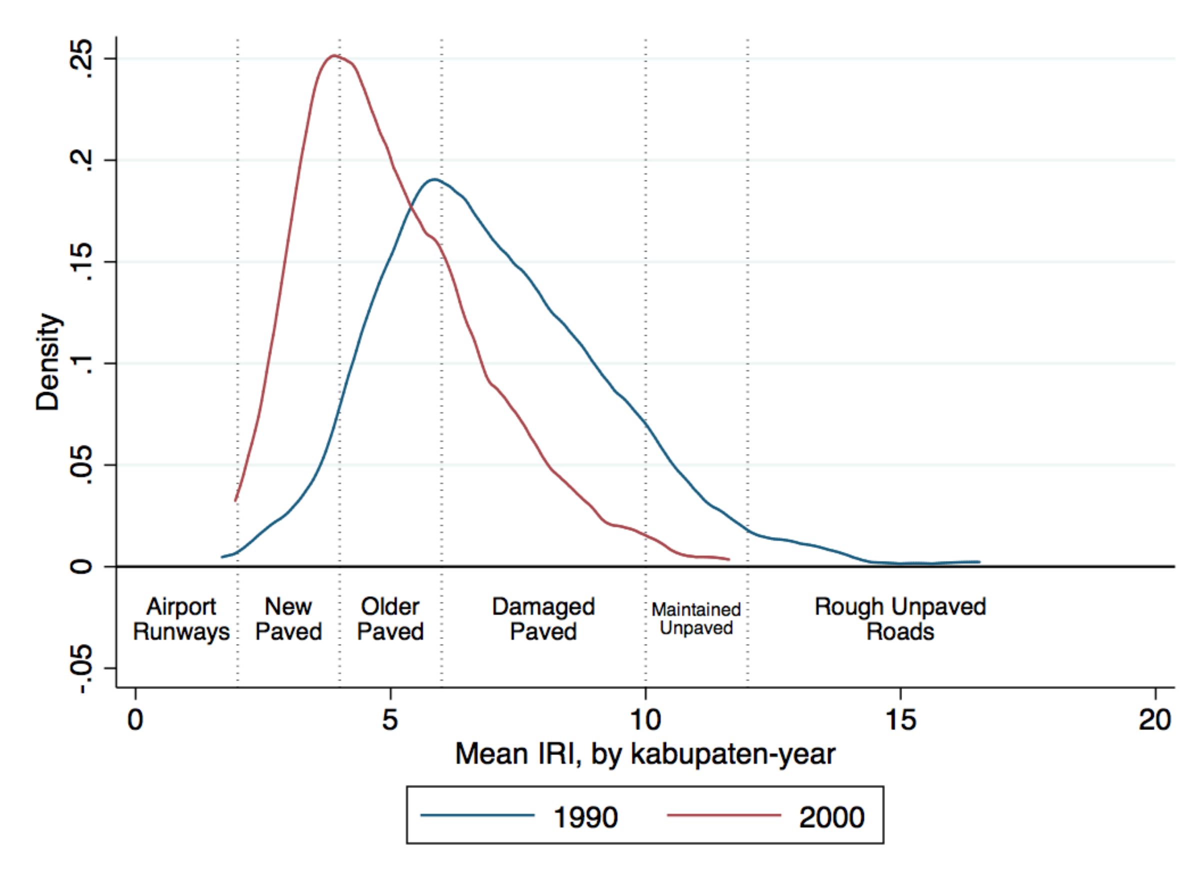
Notes: We report the results of cross-sectional regressions of the dependent variable on changes in road roughness. Each cell reports estimates of β from a separate regression, with the dependent variable listed in the row heading. For the migration regressions, controls include logs of 1990 population and 1990 non-oil GDRP. Robust standard errors reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 8: ROAD QUALITY AND PRICES

	FELS	Other District IV-Lasso		Own District IV-Lasso		Stats	
	(1)	(2)	KBP	(3)	KBP	$\overline{Y}$	N
Log Factory Wage	-0.088 (0.178)	0.358 (0.381) [0.58]	26.606 (.)	0.201 (0.374) [1.00]	21.448 (.)	3842.123 (.)	226.000 (.)
Log Farm Wage	0.050 (0.115)	-0.072 (0.172) [0.97]	128.022 (.)	0.047 (0.255) [1.00]	56.074 (.)	3766.165 (.)	339.000 (.)
Log Food Price	-0.109 (0.068)	-0.015 (0.126) [1.00]	130.677 (.)	0.050 (0.126) [1.00]	33.281 (.)	146.450 (.)	914.000 (.)
Log Perishables Price	-0.314*** (0.079)	-0.635*** (0.139) [0.02]	130.677 (.)	-0.372*** (0.139) [0.34]	33.281 (.)	76.494 (.)	914.000 (.)
Median Log Land Value	0.557** (0.254)	0.925** (0.422) [0.15]	123.408 (.)	-0.343 (0.556) [1.00]	58.778 (.)	3849.829 (.)	778.000 (.)
Median Log Rent	0.116 (0.091)	-0.182 (0.218) [0.58]	190.653 (.)	0.224 (0.161) [0.76]	32.685 (.)	3852.859 (.)	926.000 (.)
Log Land Value (Hedonic FE)	0.702*** (0.218)	1.447*** (0.390) [0.03]	160.811 (.)	0.751* (0.408) [0.61]	42.856 (.)	3785.875 (.)	622.000 (.)
Log Rent (Hedonic FE)	0.081 (0.112)	-0.194 (0.237) [0.58]	190.951 (.)	0.233 (0.204) [0.77]	37.418 (.)	3849.437 (.)	914.000 (.)

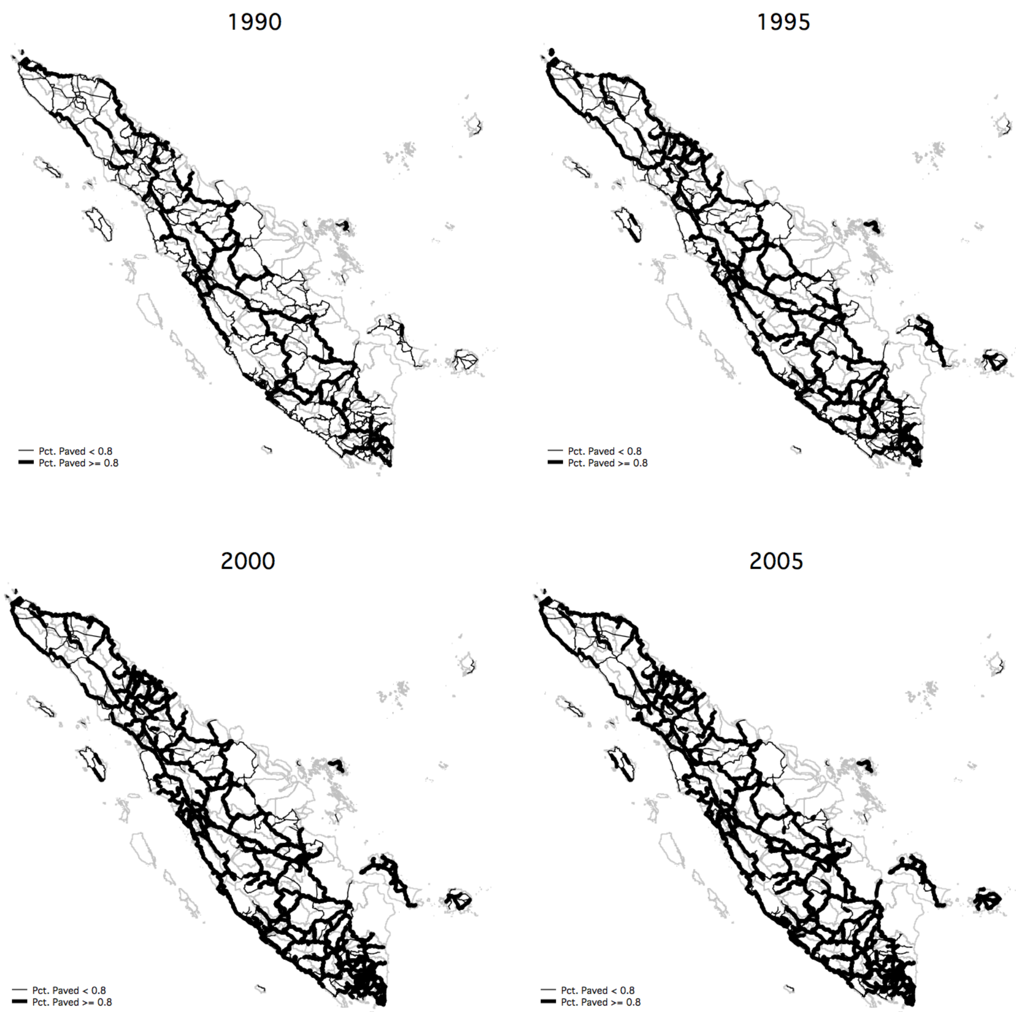
Notes: We report the results of community-level panel regressions of the dependent variable on local road quality or market potential (both in logs). Each cell reports β from a separate regression, with the dependent variable listed in the row heading. Log(Farm Wage) is not available in 1993. Dependent variable means are reported in levels. Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Figure 1: Changes in the Distribution of Road Roughness



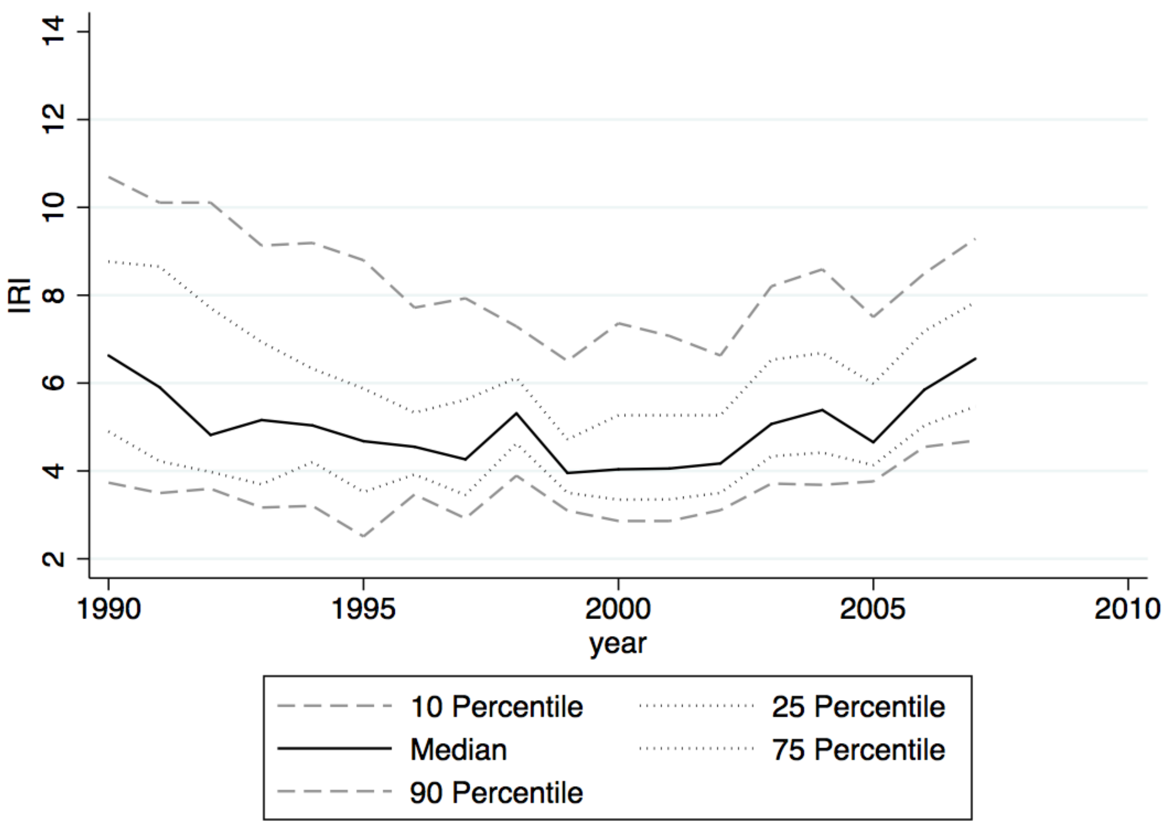
Note: Authors' calculations.

Figure 2: Road Roughness - Sumatra



Note: Authors' calculations.

Figure 3: Changes in Roughness

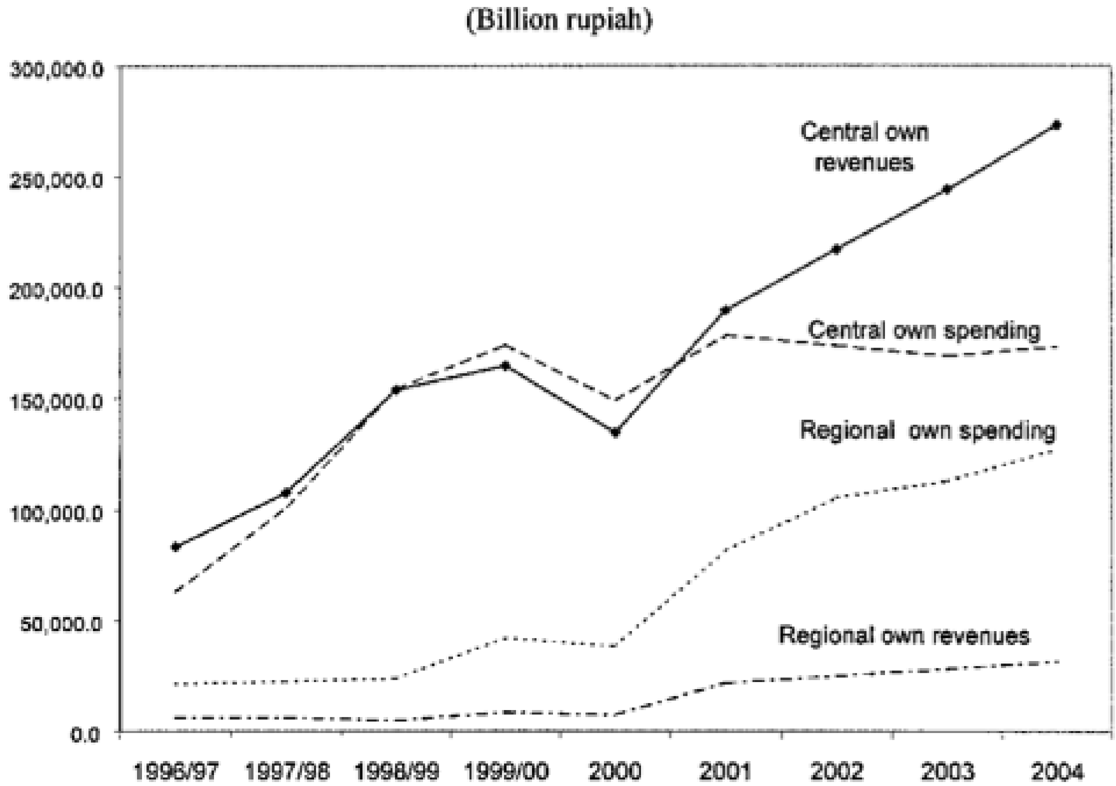


Note: Authors' calculations.

Figure 4: IFLS Villages

Note: Authors' calculations.

Figure 5: Growth in Revenues for Road Maintence



Note: From Ahmad and Mansoor (2002), in billion rupiah.

Figure 6: Allocation Criteria for District Road Improvement Grant

1989/90	1990/91	1991/92	1992/93
CRITERIA	CRITERIA	CRITERIA	CRITERIA
1. Length of road (20%) 2. Unstable & critical road (52%) 3. Area of irrigation (15%) 4. Increase of actual regional own-source receipts (PAD)(7%) 5. Actual own-source revenues compared to planned(6%) 6. Unit price correction (Ditjen Cipta Karya)	1. Length of road (15%) 2. Road condition (60%) 3. GRDP (15%) 4. Road density (10%) 5. Unit price correction	1. Length of road 2. % of good road a. Kab < 28.3% good road b. 55% > Kab > 28.3% good road 3. Road density a. Kab < 28km/1000km2 b. 100km/1000km2 > Kab > 28km/1000km2 4. Performance Kabupaten needs according to (a) and (b) greater than 60km, take 60km 5. Unit price correction	1. Length of road 2. % of good road a. Kab < 17.68% good road b. 55% > Kab > 17.68% 3. Road density a. Kab < 40.29km/1000km2 b. 100km/1000km2 kab > 40.29km/1000km2 4. Performance Kabupaten needs according to (a) and (b) greater than 60km, take 60 km 5. Unit price correction

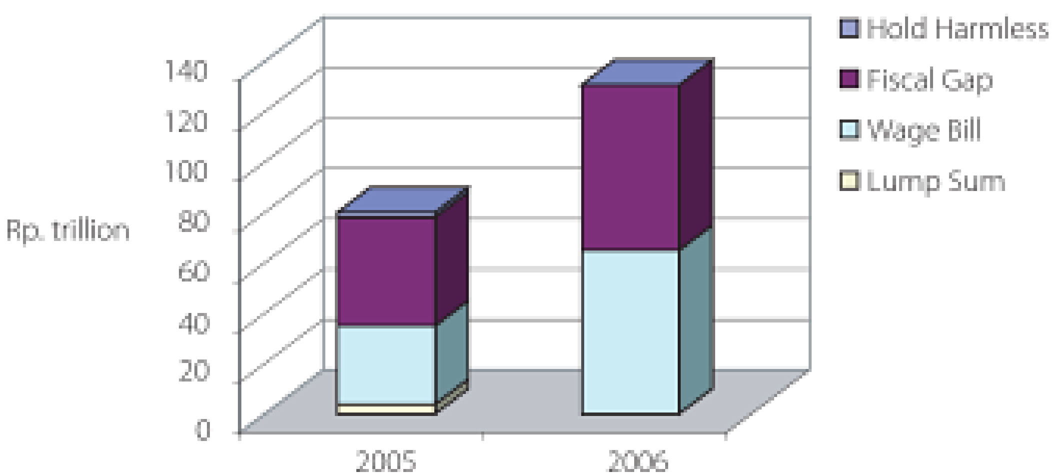
Note: Bird and Smart (2001).

Figure 7: The evolution of technical criteria in the DAK formula for roads and their respective weights

No.	Technical criteria	Description	2008	2009	2010	2011
1	Length of road	Length of road which is legally acknowledge through the decree of the head of local government	30%	25%	25%	25%
2	Road condition	Length of road with non-stable condition	30%	40%	35%	25%
3	Good road performance		20%			
4	Accessibility	Defined by the length of road divided by total area			20%	10%
5	Mobility	Length of road per 1000 population in the province/kabupaten			20%	10%
6	Ownership/concern by LG	Determined by the percentage of original APBD allocated to the road sector		20%		10%
7	Reporting	Consistency in submitting of quarterly report, physical progress, financial progress	20%	15%		20%

Note: World Bank (2012).

Figure 8: Changes in DAU composition over time



Note: World Bank staff calculation. World Bank (2008).

A Appendix Tables and Figures

Table A.1: Summary Statistics: Time-Varying Covariates

PANEL A: DISTRICT-LEVEL VARIABLES	VARNAME	MEAN	SD	N	FIRST YEAR	LAST YEAR
LOG TOTAL DAK (IDR BILLION)	DAK_TOTAL	2.33	1.24	4088	1994	2007
LOG TOTAL DBH SDA (IDR)	DBH_TOTAL	19.07	5.89	4088	1994	2007
LOG TOTAL DAU (IDR)	DAU_TOTAL	23.84	5.36	4088	1994	2007
% OF HH W/ ACCESS TO SAFE SANITATION	HHSANITATION	65.78	41.26	3504	1996	2007
% OF HH W/ ACCESS TO SAFE WATER	HHSAFEWATER	50.98	29.18	3504	1996	2007
% OF HH W/ ACCESS TO ELECTRICITY	HHELECTRICITY	82.51	51.24	3504	1996	2007
PANEL B: PROVINCE-LEVEL VARIABLES	**VARNAME**	**MEAN**	**SD**	**N**	**FIRST YEAR**	**LAST YEAR**
LOG TOTAL DAK (IDR BILLION)	DAK_TOTAL_PROV	2.33	2.10	4088	1994	2007
LOG TOTAL DBH SDA (IDR)	DBH_TOTAL_PROV	22.40	4.98	4088	1994	2007
LOG TOTAL DAU (IDR)	DAU_TOTAL_PROV	25.59	3.38	4088	1994	2007
% OF HH W/ ACCESS TO SAFE SANITATION	HHSANITATION_PROV	61.18	22.27	3504	1996	2007
% OF HH W/ ACCESS TO SAFE WATER	HHSAFEWATER_PROV	48.33	18.81	3504	1996	2007
% OF HH W/ ACCESS TO ELECTRICITY	HHELECTRICITY_PROV	79.65	41.26	3504	1996	2007
KM UPGRADED (N, PROVINCE)	KMUPGRADE_N_PROV	380.95	464.08	5117	1991	2007
KM UPGRADED (P, PROVINCE)	KMUPGRADE_P_PROV	366.42	460.26	5117	1991	2007
KM UPGRADED (N, PROVINCE, WEIGHTED)	WTKMUPGRADE_N_PROV	419.66	676.70	5117	1991	2007
KM UPGRADED (P, PROVINCE, WEIGHTED)	WTKMUPGRADE_P_PROV	602.69	949.61	5117	1991	2007
SHARE KM UPGRADED (N, PROVINCE)	SHAREKMUPGRADE_N_PROV	0.28	0.31	5117	1991	2007
SHARE KM UPGRADED (P, PROVINCE)	SHAREKMUPGRADE_P_PROV	0.24	0.27	5117	1991	2007
SHARE KM UPGRADED (N, PROVINCE, WEIGHTED)	SHAREWTKMUPGRADE_N_PROV	0.31	0.49	5117	1991	2007
SHARE KM UPGRADED (P, PROVINCE, WEIGHTED)	SHAREWTKMUPGRADE_P_PROV	0.39	0.59	5117	1991	2007

Notes: Authors' calculations.

Table A.2: Summary Statistics: Cross-Sectional Covariates

Panel A: Physical Characteristics	**Varname**	**Mean**	**Sd**	**N**
Log Area (sq meters)	area	7.39	1.76	279
Elevation (in meters), Source: HWSD	elevation	303.11	269.36	279
Percentage of Land with 0 <= slope < 0.5, Source: HWSD	slope1	0.05	0.05	279
Percentage of Land with 0.5 <= slope < 2, Source: HWSD	slope2	0.24	0.20	279
Percentage of Land with 2 <= slope < 5, Source: HWSD	slope3	0.21	0.11	279
Percentage of Land with 5 <= slope < 10, Source: HWSD	slope4	0.14	0.08	279
Percentage of Land with 10 <= slope < 15, Source: HWSD	slope5	0.08	0.05	279
Percentage of Land with 15 <= slope < 30, Source: HWSD	slope6	0.15	0.12	279
Percentage of Land with 30 <= slope < 45, Source: HWSD	slope7	0.08	0.07	279
Percentage of Land with slope >= 45, Source: HWSD	slope8	0.04	0.05	279
Average 30 arc-second Vector Ruggedness Measure (VRM)	rugged3	0.00	0.00	279
Log Avg Distance to Major Cities	d_majorcities	4.28	0.94	279
Log Avg Distance to Jakarta	d_jakarta	6.49	1.02	279
Log Avg Distance to Major Ports	d_majorports	4.28	0.82	279
Panel B: Land Cover	**Varname**	**Mean**	**Sd**	**N**
Cultivated land (percent); Source: HWSD	cult_2000	0.32	0.17	279
Forest land (percent), Source: HWSD	for_2000	0.28	0.19	279
Grass / scrub / woodland (percent), Source: HWSD	grs_2000	0.18	0.09	279
Barren / sparsely vegetated (percent), Source: HWSD	nvg_2000	0.00	0.00	279
Built-up land (percent), Source: HWSD	urb_2000	0.11	0.16	279
Water coverage (percent), Source: HWSD	wat_2000	0.03	0.07	279
Panel C: Census Data	**Varname**	**Mean**	**Sd**	**N**
Population in 1971 (millions), Census Data	pop1971	0.49	0.38	214
Population in 1980 (millions), Census Data	pop1980	0.61	0.48	214
Population in 1990 (millions), Census Data	pop1990	0.63	0.56	284
Population in 2000 (millions), Census Data	pop2000	0.72	0.66	284
Log GRDP at Current Prices ex. oil and gas, 1990	gdrp_conp_nonoil_1990	13.04	0.98	282
Panel D: Road Characteristics	**Varname**	**Mean**	**Sd**	**N**
Total Length of National Roads	tot_length_N	77.49	83.78	205
Average elevation (meters), National Roads	rd_elev_N	181.80	248.82	205
Absolute Change in Elevation, National Roads	rd_elev_change_N	31.42	41.15	205
Percent of National Roads w/ 0 <= slope < 0.5	rd_slope1_N	0.07	0.06	205
Percent of National Roads w/ 0.5 <= slope < 2	rd_slope2_N	0.38	0.21	205
Percent of National Roads w/ 2 <= slope < 5	rd_slope3_N	0.28	0.12	205
Percent of National Roads w/ 5 <= slope < 10	rd_slope4_N	0.13	0.10	205
Percent of National Roads w/ 10 <= slope < 15	rd_slope5_N	0.05	0.05	205
Percent of National Roads w/ 15 <= slope < 30	rd_slope6_N	0.06	0.08	205
Percent of National Roads w/ 30 <= slope < 45	rd_slope7_N	0.02	0.04	205
Percent of National Roads w/ slope >= 45	rd_slope8_N	0.01	0.04	205
Total Length of Provincial Roads	tot_length_P	93.05	103.63	205
Average elevation (meters), Provincial Roads	rd_elev_P	265.18	269.25	205
Absolute Change in Elevation, Provincial Roads	rd_elev_change_P	60.11	57.04	205
Percent of Provincial Roads w/ 0 <= slope < 0.5	rd_slope1_P	0.04	0.04	205
Percent of Provincial Roads w/ 0.5 <= slope < 2	rd_slope2_P	0.29	0.19	205
Percent of Provincial Roads w/ 2 <= slope < 5	rd_slope3_P	0.26	0.12	205
Percent of Provincial Roads w/ 5 <= slope < 10	rd_slope4_P	0.16	0.10	205
Percent of Provincial Roads w/ 10 <= slope < 15	rd_slope5_P	0.08	0.06	205
Percent of Provincial Roads w/ 15 <= slope < 30	rd_slope6_P	0.10	0.10	205
Percent of Provincial Roads w/ 30 <= slope < 45	rd_slope7_P	0.04	0.05	205
Percent of Provincial Roads w/ slope >= 45	rd_slope8_P	0.02	0.04	205

Notes: Authors' calculations.

Table A.3: Selected IVs for Table 3: Travel Times and Road Quality

			Hausman		Base Target	
Panel A: IFLS Travel Times	**DepVar**	**# Total**	**# Sel.**	**Selected IVs**	**# Sel.**	**Selected IVs**
Log Travel Time to Nearest District Capital	ttime_kabuC	672	3	s_pop1971_7 s_pop1971_8 s_rd_slope8_N_1	6	pop1971_7 pop1971_8 rd_slope2_P_7 rd_slope3_P_7 dbh_total_prov_12 dak_total_prov_10
Log Travel Time to Nearest Provincial Capital	ttime_provC	672	4	s_pop1971_7 s_pop1971_8 s_pop1990_8 s_rd_slope8_N_1	4	pop1971_8 rd_slope2_P_7 dbh_total_prov_12 dak_total_prov_10
Panel B: Roughness-Based Travel Times	**DepVar**	**# Total**	**# Sel.**	**Selected IVs**	**# Sel.**	**Selected IVs**
Log Travel Time to Nearest District Capital	travTimeCity	672	2	s_pop1971_8 s_rd_slope8_N_1	4	pop1971_8 rd_slope2_P_7 dbh_total_prov_12 dak_total_prov_10
Log Travel Time to Nearest Provincial Capital	travTimeProvC	672	2	s_pop1971_8 s_rd_slope8_N_1	4	pop1971_8 rd_slope2_P_7 dbh_total_prov_12 dak_total_prov_10

Notes: Notes here.

Table A.4: ROAD QUALITY AND TRAVEL TIMES (SAME PROVINCIAL CAPITALS)

	FELS	Other District IV-Lasso		Own District IV-Lasso		Stats	
	(1)	(2)	KBP	(3)	KBP	$\overline{Y}$	N
Log Travel Time to Nearest Prov. Capital (IFLS)	-0.185**	-0.266*	100.027	-0.427**	30.907	169.363	856.000
	(0.074)	(0.141)	(.)	(0.208)	(.)	(.)	(.)
		[0.06]		[0.05]			
Log Travel Time to Prov. Capital (Roughness-Based)	-0.454***	-0.994***	105.253	-1.027***	44.463	73.002	888.000
	(0.031)	(0.091)	(.)	(0.097)	(.)	(.)	(.)
		[0.00]		[0.00]			
Year FE	Yes	Yes		Yes			
Village FE	Yes	Yes		Yes			

Notes: We report the results of community-level panel regressions of the dependent variable on road quality (both in logs). Each cell reports β from a separate regression, with the dependent variable listed in the row heading. All regressions include community and year fixed effects, with controls including logs of current population and non-oil GDRP. Interpretation of results remains unchanged when dependent variables are expressed in levels. Dependent variable means are reported in levels. Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.5: ISLAND MARKET POTENTIAL AND TRAVEL TIMES

	FELS	Other District IV-Lasso		Own District IV-Lasso		Stats	
Panel A: IFLS Travel Times	(1)	(2)	KBP	(3)	KBP	$\overline{Y}$	N
Log Travel Time to Nearest District Capital	-0.252	-0.143	292.226	-0.525*	77.778	46.241	850.000
	(0.207)	(0.193)	(.)	(0.303)	(.)	(.)	(.)
		[0.32]		[0.07]			
Log Travel Time to Nearest Provincial Capital	-0.147	-0.181	427.972	-0.360*	77.927	172.334	872.000
	(0.107)	(0.132)	(.)	(0.201)	(.)	(.)	(.)
		[0.17]		[0.07]			
Panel B: Roughness-Based Travel Times							
Log Travel Time to Nearest District Capital	-0.815***	-0.546***	477.431	-0.576***	83.620	21.272	904.000
	(0.125)	(0.087)	(.)	(0.110)	(.)	(.)	(.)
		[0.00]		[0.00]			
Log Travel Time to Nearest Provincial Capital	-0.890***	-0.951***	477.431	-0.929***	83.620	75.217	904.000
	(0.042)	(0.051)	(.)	(0.068)	(.)	(.)	(.)
		[0.00]		[0.00]			

Notes: We report the results of community-level panel regressions of the dependent variable on market potential (both in logs). Each cell reports β from a separate regression, with the dependent variable listed in the row heading. All regressions include community and year fixed effects, with controls including logs of current population and non-oil GDRP. Interpretation of results remains unchanged when dependent variables are expressed in levels. Dependent variable means are reported in levels. Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.6: ROAD QUALITY AND TRAVEL TIMES: ROBUSTNESS, HAUSMAN IV

		IV-Lasso							
	FELS	**Hausman**	**No Oil / Gas**	**No Mining**	**PLN**	**TV**	**Schools**	**Health Facil.**	**Only 1998+**
Panel A: IFLS Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Log Travel Time to Nearest District Capital	-0.092	-0.346	-0.215	-0.612	-0.471	-0.435	-0.283	-0.248	0.250
	(0.145)	(0.298)	(0.359)	(0.376)	(0.287)	(0.286)	(0.295)	(0.321)	(0.367)
N	850	850	722	630	850	850	850	850	410
Kleibergen-Paap Wald Rank *F* Stat		76.90	63.31	31.68	75.69	70.73	76.35	68.16	63.47
Log Travel Time to Nearest Provincial Capital	-0.203***	-0.221	-0.127	0.125	-0.173	-0.198	-0.217	-0.207	-0.007
	(0.073)	(0.138)	(0.167)	(0.338)	(0.139)	(0.147)	(0.135)	(0.146)	(0.255)
N	872	872	732	650	872	872	872	872	416
Kleibergen-Paap Wald Rank *F* Stat		62.47	65.33	33.51	61.51	55.48	64.52	56.18	54.79
Panel B: Roughness-Based Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Log Travel Time to Nearest District Capital	-0.487***	-0.731***	-0.790***	-0.275**	-0.706***	-0.720***	-0.701***	-0.732***	-0.986***
	(0.049)	(0.107)	(0.112)	(0.118)	(0.103)	(0.107)	(0.113)	(0.111)	(0.185)
N	904	904	760	672	904	904	904	904	454
Kleibergen-Paap Wald Rank *F* Stat		112.42	69.47	37.56	109.64	99.18	113.47	99.13	70.19
Log Travel Time to Nearest Provincial Capital	-0.447***	-0.968***	-0.977***	-0.419***	-0.946***	-1.007***	-0.955***	-1.019***	-1.036***
	(0.032)	(0.094)	(0.101)	(0.091)	(0.093)	(0.097)	(0.092)	(0.103)	(0.103)
N	904	904	760	672	904	904	904	904	454
Kleibergen-Paap Wald Rank *F* Stat		112.42	69.47	37.56	109.64	99.18	113.47	99.13	70.19

Notes: We report the results of community-level panel regressions of the dependent variable on market potential (both in logs). Each cell reports β from a separate regression, with the dependent variable listed in the row heading. All regressions include community and year fixed effects, with controls including logs of current population and non-oil GDRP. Interpretation of results remains unchanged when dependent variables are expressed in levels. Dependent variable means are reported in levels. Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.7: ROAD QUALITY AND TRAVEL TIMES: ROBUSTNESS, BTARGET IV

		IV-Lasso							
	FELS	**bTarget**	**No Oil / Gas**	**No Mining**	**PLN**	**TV**	**Schools**	**Health Facil.**	**Only 1998+**
Panel A: IFLS Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Log Travel Time to Nearest District Capital	-0.092	-0.733**	-0.136	-0.612	-0.811**	-0.656**	-0.789**	-0.708*	0.017
	(0.145)	(0.326)	(0.422)	(0.376)	(0.330)	(0.320)	(0.340)	(0.392)	(0.310)
N	850	850	722	630	850	850	850	850	410
Kleibergen-Paap Wald Rank *F* Stat		22.60	28.76	31.68	25.04	26.72	23.19	20.98	115.97
Log Travel Time to Nearest Provincial Capital	-0.203***	-0.440**	-0.172	0.125	-0.433**	-0.420**	-0.498**	-0.461*	0.520*
	(0.073)	(0.212)	(0.249)	(0.338)	(0.211)	(0.206)	(0.213)	(0.235)	(0.273)
N	872	872	732	650	872	872	872	872	416
Kleibergen-Paap Wald Rank *F* Stat		33.60	31.84	33.51	33.49	36.31	32.95	28.32	118.86
Panel B: Roughness-Based Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Log Travel Time to Nearest District Capital	-0.487***	-0.503***	-0.391***	-0.275**	-0.508***	-0.525***	-0.430***	-0.371***	-0.666***
	(0.049)	(0.101)	(0.106)	(0.118)	(0.099)	(0.097)	(0.099)	(0.114)	(0.205)
N	904	904	760	672	904	904	904	904	454
Kleibergen-Paap Wald Rank *F* Stat		38.41	35.23	37.56	37.68	41.07	37.53	32.21	117.75
Log Travel Time to Nearest Provincial Capital	-0.447***	-1.016***	-0.795***	-0.419***	-1.033***	-0.973***	-1.006***	-1.133***	-0.593***
	(0.032)	(0.098)	(0.098)	(0.091)	(0.098)	(0.093)	(0.099)	(0.114)	(0.107)
N	904	904	760	672	904	904	904	904	454
Kleibergen-Paap Wald Rank *F* Stat		38.41	35.23	37.56	37.68	41.07	37.53	32.21	117.75

Notes: We report the results of community-level panel regressions of the dependent variable on market potential (both in logs). Each cell reports β from a separate regression, with the dependent variable listed in the row heading. All regressions include community and year fixed effects, with controls including logs of current population and non-oil GDRP. Interpretation of results remains unchanged when dependent variables are expressed in levels. Dependent variable means are reported in levels. Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.8: Road Quality and Travel Times: Heterogeneity, Hausman IV

		IV-Lasso				
			GHSL 1990		Pop 1990	
	FELS	**Hausman**	**$<$ 20 pct.**	**$\geq$ 20 pct.**	**$<$ 1 mil.**	**$\geq$ 1 mil.**
Panel A: IFLS Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Log Travel Time to Nearest District Capital	-0.092	-0.346	-0.386	0.463	-0.151	-0.437
	(0.145)	(0.298)	(0.250)	(1.143)	(0.398)	(0.333)
N	850	850	611	239	511	339
Kleibergen-Paap Wald Rank *F* Stat		76.90	71.49	9.22	45.99	58.96
Log Travel Time to Nearest Provincial Capital	-0.203***	-0.221	-0.273*	0.266	-0.070	-0.341
	(0.073)	(0.138)	(0.152)	(0.315)	(0.174)	(0.253)
N	872	872	627	245	526	346
Kleibergen-Paap Wald Rank *F* Stat		62.47	55.96	15.14	39.00	50.57
Panel B: Roughness-Based Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Log Travel Time to Nearest District Capital	-0.487***	-0.731***	-0.778***	-1.273*	-0.820***	-0.933***
	(0.049)	(0.107)	(0.122)	(0.725)	(0.218)	(0.160)
N	904	904	647	257	544	360
Kleibergen-Paap Wald Rank *F* Stat		112.42	102.62	16.94	65.02	90.43
Log Travel Time to Nearest Provincial Capital	-0.447***	-0.968***	-0.995***	-0.736***	-0.944***	-0.957***
	(0.032)	(0.094)	(0.099)	(0.214)	(0.127)	(0.107)
N	904	904	647	257	544	360
Kleibergen-Paap Wald Rank *F* Stat		112.42	102.62	16.94	65.02	90.43

Notes: Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.9: Road Quality and Travel Times: Heterogeneity, bTarget IV

		IV-Lasso				
			GHSL 1990		Pop 1990	
	FELS	**bTarget**	**$<$ 20 pct.**	**$\geq$ 20 pct.**	**$<$ 1 mil.**	**$\geq$ 1 mil.**
Panel A: IFLS Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Log Travel Time to Nearest District Capital	-0.092	-0.733**	-0.541*	-0.715	-0.607	-0.329
	(0.145)	(0.326)	(0.292)	(0.634)	(0.413)	(0.395)
N	850	850	611	239	511	339
Kleibergen-Paap Wald Rank F Stat		22.60	18.59	27.53	14.68	15.18
Log Travel Time to Nearest Provincial Capital	-0.203***	-0.440**	-0.580***	-0.423	-0.702***	-0.043
	(0.073)	(0.212)	(0.215)	(0.739)	(0.267)	(0.338)
N	872	872	627	245	526	346
Kleibergen-Paap Wald Rank F Stat		33.60	28.28	17.80	17.24	15.09
Panel B: Roughness-Based Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Log Travel Time to Nearest District Capital	-0.487***	-0.503***	-0.516***	-0.279**	-0.384***	-0.566***
	(0.049)	(0.101)	(0.112)	(0.138)	(0.131)	(0.151)
N	904	904	647	257	544	360
Kleibergen-Paap Wald Rank F Stat		38.41	28.97	21.07	19.83	16.55
Log Travel Time to Nearest Provincial Capital	-0.447***	-1.016***	-0.996***	-0.203**	-0.967***	-0.584***
	(0.032)	(0.098)	(0.108)	(0.092)	(0.131)	(0.109)
N	904	904	647	257	544	360
Kleibergen-Paap Wald Rank F Stat		38.41	28.97	21.07	19.83	16.55

Notes: Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.10: Selected IVs for Table 4: Road Quality and District-Level Manufacturing Outcomes

			Hausman		Base Target	
Panel A: Firm Counts	**DepVar**	**# Total**	**# Sel.**	**Selected IVs**	**# Sel.**	**Selected IVs**
Any Firms (0 1)	anyFirms	448	6	s_pop1990_4 s_wat_2000_4 s_slope2_4 s_rd_slope3_P_4 s_dak_total_4 s_dak_total_prov_4	6	pop1971_3 pop1971_4 rd_slope2_N_4 rd_slope2_P_3 dak_total_4 dau_total_prov_4
Log Number of Opened Firms	logOpen	448	6	s_pop1990_4 s_wat_2000_4 s_slope2_4 s_rd_slope3_P_4 s_dak_total_4 s_dak_total_prov_4	6	pop1971_3 pop1971_4 rd_slope2_N_4 rd_slope2_P_3 dak_total_4 dau_total_prov_4
Log Number of Closed Firms	logClose	448	6	s_pop1990_4 s_wat_2000_4 s_rd_slope3_N_4 s_tot_length_N_4 s_dak_total_4 s_dak_total_prov_4	6	pop1971_3 pop1971_4 d_majorcities_4 gdrp_conp_nonoil_1990_3 dak_total_4 dau_total_prov_4
Percent Δ Number of Firms	pctDeltaFirms	448	6	s_pop1990_4 s_wat_2000_4 s_slope2_4 s_rd_slope3_P_4 s_dak_total_4 s_dak_total_prov_4	8	pop1971_3 pop1971_4 d_majorcities_4 rd_slope2_N_3 rd_slope2_N_4 rd_slope2_P_3 dak_total_4 dau_total_prov_4
Panel B: Production	**DepVar**	**# Total**	**# Sel.**	**Selected IVs**	**# Sel.**	**Selected IVs**
Log Output	logOutput	448	6	s_pop1990_4 s_wat_2000_4 s_slope2_4 s_rd_slope3_P_4 s_dak_total_4 s_dak_total_prov_4	6	pop1971_3 pop1971_4 rd_slope2_N_4 rd_slope2_P_3 dak_total_4 dau_total_prov_4
Log Value Added	logValAdded	448	6	s_pop1990_4 s_wat_2000_4 s_slope2_4 s_rd_slope3_P_4 s_dak_total_4 s_dak_total_prov_4	6	pop1971_3 pop1971_4 rd_slope2_N_4 rd_slope2_P_3 dak_total_4 dau_total_prov_4
Log Total Employment	logNumWorkers	448	6	s_pop1990_4 s_wat_2000_4 s_slope2_4 s_rd_slope3_P_4 s_dak_total_4 s_dak_total_prov_4	6	pop1971_3 pop1971_4 rd_slope2_N_4 rd_slope2_P_3 dak_total_4 dau_total_prov_4
Log Output per Worker	logOutputPerWorker	448	6	s_pop1990_4 s_wat_2000_4 s_slope2_4 s_rd_slope3_P_4 s_dak_total_4 s_dak_total_prov_4	6	pop1971_3 pop1971_4 rd_slope2_N_4 rd_slope2_P_3 dak_total_4 dau_total_prov_4

Notes: Notes here.

Table A.11: Island Market Potential and District-Level Manufacturing Outcomes

	FELS	Other District IV-Lasso		Own District IV-Lasso		Stats	
Panel A: Firm Counts	**(1)**	**(2)**	**KBP**	**(3)**	**KBP**	$\overline{Y}$	N
Any Firms (0 1)	0.130***	0.091*	107.823	0.112**	78.434	0.960	3449.000
	(0.040)	(0.051)	(.)	(0.055)	(.)	(.)	(.)
		[0.22]		[0.11]			
Log Number of Opened Firms	0.414***	0.962***	107.823	0.730***	78.434	6.918	3449.000
	(0.125)	(0.248)	(.)	(0.236)	(.)	(.)	(.)
		[0.01]		[0.05]			
Log Number of Closed Firms	-0.213	-0.149	273.154	-0.692**	108.491	6.762	3248.000
	(0.152)	(0.365)	(.)	(0.285)	(.)	(.)	(.)
		[0.75]		[0.10]			
Percent Δ Number of Firms	0.128***	0.176***	107.554	0.151**	77.204	-0.030	3409.000
	(0.046)	(0.066)	(.)	(0.070)	(.)	(.)	(.)
		[0.06]		[0.11]			
Panel B: Production							
Log Output	1.035***	0.493	107.823	0.509	78.434	2032.332	3449.000
	(0.376)	(0.468)	(.)	(0.428)	(.)	(.)	(.)
		[0.44]		[0.25]			
Log Value Added	0.807**	0.241	107.823	0.188	78.434	745.998	3449.000
	(0.356)	(0.487)	(.)	(0.442)	(.)	(.)	(.)
		[0.75]		[0.52]			
Log Total Employment	0.049	0.106	107.823	0.220	78.434	16931.914	3449.000
	(0.184)	(0.244)	(.)	(0.206)	(.)	(.)	(.)
		[0.75]		[0.25]			
Log Output per Worker	0.492***	0.316	80.402	0.095	62.155	0.084	3309.000
	(0.167)	(0.269)	(.)	(0.231)	(.)	(.)	(.)
		[0.44]		[0.52]			
Year FE	Yes	Yes		Yes			
District FE	Yes	Yes		Yes			

Notes: We report the results of district-level panel regressions of the dependent variable on island market potential. Each cell reports β from a separate regression, with the dependent variable listed in the row heading. All regressions include district and year fixed effects, with controls that include logs of current population and non-oil GDRP. Robust standard errors, clustered at the district level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.12: Road Quality and District-Level Manufacturing Outcomes: Firm Counts (Robustness, Hausman IV)

		IV-Lasso							
	FELS	**Hausman**	**No Oil / Gas**	**No Mining**	**PLN**	**TV**	**Schools**	**Health Facil.**	**Only 1998+**
Panel A: IFLS Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Any Firms (0 1)	0.007	0.073*	0.106**	0.161***	0.071*	0.063*	0.071*	0.063	0.104
	(0.017)	(0.039)	(0.053)	(0.060)	(0.039)	(0.037)	(0.037)	(0.039)	(0.075)
N	3381	3381	2803	2480	3377	3377	3377	3377	1789
Kleibergen-Paap Wald Rank F Stat		69.02	58.56	69.88	63.85	64.26	67.61	70.19	132.37
Log Number of Opened Firms	0.058	0.602***	0.643***	0.503***	0.615***	0.605***	0.561***	0.605***	0.926***
	(0.060)	(0.134)	(0.158)	(0.154)	(0.136)	(0.139)	(0.135)	(0.136)	(0.353)
N	3381	3381	2803	2480	3377	3377	3377	3377	1789
Kleibergen-Paap Wald Rank F Stat		69.02	58.56	69.88	63.85	64.26	67.61	70.19	132.37
Log Number of Closed Firms	-0.042	-0.089	0.332	0.056	-0.104	-0.087	-0.147	-0.091	1.125**
	(0.070)	(0.194)	(0.229)	(0.228)	(0.195)	(0.194)	(0.192)	(0.194)	(0.519)
N	3184	3184	2640	2336	3180	3180	3180	3180	1592
Kleibergen-Paap Wald Rank F Stat		71.13	105.00	62.56	68.84	68.64	70.23	71.78	57.70
Percent Δ Number of Firms	-0.003	0.120***	0.133**	0.214***	0.129***	0.133***	0.114***	0.119***	0.086
	(0.020)	(0.044)	(0.056)	(0.077)	(0.045)	(0.047)	(0.044)	(0.045)	(0.096)
N	3337	3337	2759	2436	3335	3335	3335	3335	1759
Kleibergen-Paap Wald Rank F Stat		69.36	58.31	91.15	64.28	64.61	67.92	70.79	130.16

Notes: We report the results of community-level panel regressions of the dependent variable on market potential (both in logs). Each cell reports β from a separate regression, with the dependent variable listed in the row heading. All regressions include community and year fixed effects, with controls including logs of current population and non-oil GDRP. Interpretation of results remains unchanged when dependent variables are expressed in levels. Dependent variable means are reported in levels. Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.13: Road Quality and District-Level Manufacturing Outcomes: Firm Counts (Robustness, BTarget IV)

		IV-Lasso							
	FELS	**BTarget**	**No Oil / Gas**	**No Mining**	**PLN**	**TV**	**Schools**	**Health Facil.**	**Only 1998+**
Panel A: IFLS Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Any Firms (0 1)	0.007	0.086**	0.125***	0.161***	0.061*	0.070**	0.068**	0.063*	0.057
	(0.017)	(0.034)	(0.042)	(0.060)	(0.034)	(0.033)	(0.033)	(0.034)	(0.073)
N	3381	3381	2803	2480	3377	3377	3377	3377	1789
Kleibergen-Paap Wald Rank F Stat		87.20	58.17	69.88	74.23	70.63	75.05	78.64	62.09
Log Number of Opened Firms	0.058	0.589***	0.557***	0.503***	0.619***	0.609***	0.563***	0.596***	0.834**
	(0.060)	(0.139)	(0.149)	(0.154)	(0.142)	(0.141)	(0.140)	(0.140)	(0.394)
N	3381	3381	2803	2480	3377	3377	3377	3377	1789
Kleibergen-Paap Wald Rank F Stat		87.20	58.17	69.88	74.23	70.63	75.05	78.64	62.09
Log Number of Closed Firms	-0.042	-0.331*	-0.171	0.056	-0.319*	-0.318*	-0.362**	-0.328*	0.359
	(0.070)	(0.184)	(0.205)	(0.228)	(0.186)	(0.184)	(0.184)	(0.182)	(0.387)
N	3184	3184	2640	2336	3180	3180	3180	3180	1592
Kleibergen-Paap Wald Rank F Stat		91.13	73.12	62.56	87.59	83.66	89.65	92.95	63.94
Percent Δ Number of Firms	-0.003	0.161***	0.185***	0.214***	0.159***	0.158***	0.139***	0.157***	0.200
	(0.020)	(0.042)	(0.056)	(0.077)	(0.043)	(0.046)	(0.041)	(0.043)	(0.122)
N	3337	3337	2759	2436	3335	3335	3335	3335	1759
Kleibergen-Paap Wald Rank F Stat		69.71	69.03	91.15	68.73	72.02	68.81	72.95	61.97

Notes: We report the results of community-level panel regressions of the dependent variable on market potential (both in logs). Each cell reports β from a separate regression, with the dependent variable listed in the row heading. All regressions include community and year fixed effects, with controls including logs of current population and non-oil GDRP. Interpretation of results remains unchanged when dependent variables are expressed in levels. Dependent variable means are reported in levels. Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.14: Road Quality and District-Level Manufacturing Outcomes: Production (Robustness, Hausman IV)

		IV-Lasso							
	FELS	**Hausman**	**No Oil / Gas**	**No Mining**	**PLN**	**TV**	**Schools**	**Health Facil.**	**Only 1998+**
Panel A: IFLS Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Log Output	0.118	0.526**	0.625**	0.926***	0.553**	0.572**	0.460*	0.482**	0.798
	(0.144)	(0.244)	(0.317)	(0.325)	(0.247)	(0.251)	(0.242)	(0.244)	(0.605)
N	3381	3381	2803	2480	3377	3377	3377	3377	1789
Kleibergen-Paap Wald Rank F Stat		69.02	58.56	69.88	63.85	64.26	67.61	70.19	132.37
Log Value Added	0.078	0.434*	0.552*	0.734**	0.444*	0.493**	0.374	0.402*	0.562
	(0.140)	(0.244)	(0.311)	(0.317)	(0.245)	(0.247)	(0.242)	(0.243)	(0.607)
N	3381	3381	2803	2480	3377	3377	3377	3377	1789
Kleibergen-Paap Wald Rank F Stat		69.02	58.56	69.88	63.85	64.26	67.61	70.19	132.37
Log Total Employment	-0.169**	0.253**	0.250*	0.243	0.265**	0.287**	0.229*	0.238**	0.411
	(0.081)	(0.120)	(0.150)	(0.160)	(0.122)	(0.121)	(0.121)	(0.120)	(0.251)
N	3381	3381	2803	2480	3377	3377	3377	3377	1789
Kleibergen-Paap Wald Rank F Stat		69.02	58.56	69.88	63.85	64.26	67.61	70.19	132.37
Log Output per Worker	0.296***	0.304**	0.291*	0.351**	0.310**	0.308**	0.265**	0.302**	0.178
	(0.086)	(0.136)	(0.159)	(0.153)	(0.136)	(0.136)	(0.134)	(0.135)	(0.327)
N	3232	3232	2657	2332	3230	3230	3230	3230	1740
Kleibergen-Paap Wald Rank F Stat		67.04	56.53	53.20	61.84	62.31	65.54	78.95	75.12

Notes: We report the results of community-level panel regressions of the dependent variable on market potential (both in logs). Each cell reports β from a separate regression, with the dependent variable listed in the row heading. All regressions include community and year fixed effects, with controls including logs of current population and non-oil GDRP. Interpretation of results remains unchanged when dependent variables are expressed in levels. Dependent variable means are reported in levels. Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.15: ROAD QUALITY AND DISTRICT-LEVEL MANUFACTURING OUTCOMES: PRODUCTION (ROBUSTNESS, BTARGET)

		IV-Lasso							
	FELS	**bTarget**	**No Oil / Gas**	**No Mining**	**PLN**	**TV**	**Schools**	**Health Facil.**	**Only 1998+**
Panel A: IFLS Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Log Output	0.118	0.805***	0.871***	0.926***	0.744***	0.665***	0.638***	0.709***	0.622
	(0.144)	(0.249)	(0.275)	(0.325)	(0.251)	(0.241)	(0.240)	(0.244)	(0.570)
N	3381	3381	2803	2480	3377	3377	3377	3377	1789
Kleibergen-Paap Wald Rank *F* Stat		87.20	58.17	69.88	74.23	70.63	75.05	78.64	62.09
Log Value Added	0.078	0.705***	0.718***	0.734**	0.613**	0.605**	0.519**	0.585**	0.465
	(0.140)	(0.254)	(0.274)	(0.317)	(0.250)	(0.248)	(0.244)	(0.246)	(0.591)
N	3381	3381	2803	2480	3377	3377	3377	3377	1789
Kleibergen-Paap Wald Rank *F* Stat		87.20	58.17	69.88	74.23	70.63	75.05	78.64	62.09
Log Total Employment	-0.169**	0.271**	0.252*	0.243	0.247*	0.278**	0.176	0.165	0.638**
	(0.081)	(0.130)	(0.147)	(0.160)	(0.128)	(0.127)	(0.127)	(0.123)	(0.255)
N	3381	3381	2803	2480	3377	3377	3377	3377	1789
Kleibergen-Paap Wald Rank *F* Stat		87.20	58.17	69.88	74.23	70.63	75.05	78.64	62.09
Log Output per Worker	0.296***	0.351***	0.445***	0.351**	0.365***	0.385***	0.344***	0.347***	-0.160
	(0.086)	(0.130)	(0.137)	(0.153)	(0.129)	(0.132)	(0.128)	(0.130)	(0.345)
N	3232	3232	2657	2332	3230	3230	3230	3230	1740
Kleibergen-Paap Wald Rank *F* Stat		87.30	63.42	53.20	85.15	68.02	86.76	92.33	65.72

Notes: We report the results of community-level panel regressions of the dependent variable on market potential (both in logs). Each cell reports β from a separate regression, with the dependent variable listed in the row heading. All regressions include community and year fixed effects, with controls including logs of current population and non-oil GDRP. Interpretation of results remains unchanged when dependent variables are expressed in levels. Dependent variable means are reported in levels. Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.16: ROAD QUALITY AND DISTRICT-LEVEL MANUFACTURING OUTCOMES: HETEROGENEITY, HAUSMAN IV

		IV-Lasso				
			GHSL 1990		Pop 1990	
	FELS	**Hausman**	**< 20 pct.**	**≥ 20 pct.**	**< 1 mil.**	**≥ 1 mil.**
Panel A: Firm Counts	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Any Firms (0 1)	0.007	0.073*	0.031	0.100*	0.073	-0.002
	(0.017)	(0.039)	(0.062)	(0.054)	(0.084)	(0.002)
N	3381	3381	1393	1988	1393	1988
Kleibergen-Paap Wald Rank *F* Stat		69.02	64.76	36.72	21.12	46.91
Log Number of Opened Firms	0.058	0.602***	-0.049	0.639***	0.795***	0.519**
	(0.060)	(0.134)	(0.198)	(0.178)	(0.235)	(0.208)
N	3381	3381	1393	1988	1393	1988
Kleibergen-Paap Wald Rank *F* Stat		69.02	64.76	36.72	21.12	46.91
Log Number of Closed Firms	-0.042	-0.089	-0.044	0.421	-0.163	0.003
	(0.070)	(0.194)	(0.271)	(0.305)	(0.307)	(0.294)
N	3184	3184	1312	1872	1312	1872
Kleibergen-Paap Wald Rank *F* Stat		71.13	53.32	35.63	23.55	38.43
Percent Δ Number of Firms	-0.003	0.120***	0.048	0.121*	0.090	0.022
	(0.020)	(0.044)	(0.084)	(0.066)	(0.105)	(0.029)
N	3337	3337	1349	1988	1349	1988
Kleibergen-Paap Wald Rank *F* Stat		69.36	65.41	36.72	21.14	46.91
Panel B: Production						
Log Output	0.118	0.526**	0.014	0.092	0.949**	-0.086
	(0.144)	(0.244)	(0.489)	(0.359)	(0.441)	(0.329)
N	3381	3381	1393	1988	1393	1988
Kleibergen-Paap Wald Rank *F* Stat		69.02	64.76	36.72	21.12	46.91
Log Value Added	0.078	0.434*	-0.009	0.095	0.943**	-0.170
	(0.140)	(0.244)	(0.471)	(0.357)	(0.429)	(0.367)
N	3381	3381	1393	1988	1393	1988
Kleibergen-Paap Wald Rank *F* Stat		69.02	64.76	36.72	21.12	46.91
Log Total Employment	-0.169**	0.253**	-0.107	-0.135	0.306	0.026
	(0.081)	(0.120)	(0.275)	(0.174)	(0.202)	(0.189)
N	3381	3381	1393	1988	1393	1988
Kleibergen-Paap Wald Rank *F* Stat		69.02	64.76	36.72	21.12	46.91
Log Output per Worker	0.296***	0.304**	0.347	0.019	0.480*	0.024
	(0.086)	(0.136)	(0.303)	(0.149)	(0.279)	(0.217)
N	3232	3232	1267	1965	1245	1987
Kleibergen-Paap Wald Rank *F* Stat		67.04	62.33	34.47	18.92	46.86

Notes: Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.17: ROAD QUALITY AND DISTRICT-LEVEL MANUFACTURING OUTCOMES: HETEROGENEITY, BTARGET IV

		IV-Lasso				
			GHSL 1990		Pop 1990	
	FELS	bTarget	$<$ 20 pct.	$\geq$ 20 pct.	$<$ 1 mil.	$\geq$ 1 mil.
Panel A: Firm Counts	(1)	(2)	(3)	(4)	(5)	(6)
Any Firms (0 1)	0.007	0.086**	0.147*	0.083*	0.178	0.004
	(0.017)	(0.034)	(0.085)	(0.050)	(0.110)	(0.004)
N	3381	3381	1393	1988	1393	1988
Kleibergen-Paap Wald Rank F Stat		87.20	35.19	45.21	17.37	55.95
Log Number of Opened Firms	0.058	0.589***	0.073	0.635***	0.987***	0.405*
	(0.060)	(0.139)	(0.256)	(0.181)	(0.268)	(0.231)
N	3381	3381	1393	1988	1393	1988
Kleibergen-Paap Wald Rank F Stat		87.20	35.19	45.21	17.37	55.95
Log Number of Closed Firms	-0.042	-0.331*	-0.322	0.318	-0.362	-0.575**
	(0.070)	(0.184)	(0.308)	(0.281)	(0.327)	(0.281)
N	3184	3184	1312	1872	1312	1872
Kleibergen-Paap Wald Rank F Stat		91.13	38.28	42.86	23.51	51.91
Percent Δ Number of Firms	-0.003	0.161***	0.288**	0.083	0.243**	0.018
	(0.020)	(0.042)	(0.114)	(0.067)	(0.124)	(0.029)
N	3337	3337	1349	1988	1349	1988
Kleibergen-Paap Wald Rank F Stat		69.71	31.32	35.61	15.62	42.76
Panel B: Production						
Log Output	0.118	0.805***	0.971*	0.090	1.805***	0.023
	(0.144)	(0.249)	(0.544)	(0.340)	(0.576)	(0.219)
N	3381	3381	1393	1988	1393	1988
Kleibergen-Paap Wald Rank F Stat		87.20	35.19	45.21	17.37	55.95
Log Value Added	0.078	0.705***	0.812	-0.059	1.506***	-0.083
	(0.140)	(0.254)	(0.509)	(0.341)	(0.546)	(0.255)
N	3381	3381	1393	1988	1393	1988
Kleibergen-Paap Wald Rank F Stat		87.20	35.19	45.21	17.37	55.95
Log Total Employment	-0.169**	0.271**	0.278	0.026	0.869***	0.080
	(0.081)	(0.130)	(0.251)	(0.163)	(0.283)	(0.112)
N	3381	3381	1393	1988	1393	1988
Kleibergen-Paap Wald Rank F Stat		87.20	35.19	45.21	17.37	55.95
Log Output per Worker	0.296***	0.351***	0.339	-0.022	0.365	0.003
	(0.086)	(0.130)	(0.272)	(0.140)	(0.297)	(0.131)
N	3232	3232	1267	1965	1245	1987
Kleibergen-Paap Wald Rank F Stat		87.30	37.96	42.10	15.86	56.03

Notes: Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.18: SELECTED IVS FOR TABLE 5: ROAD QUALITY AND FIRM-LEVEL MANUFACTURING OUTCOMES

	DepVar	# Total	Hausman		Base Target	
			# Sel.	Selected IVs	# Sel.	Selected IVs
Log Output	logOutput	448	3	s_pop1990_4 s_dak_total_4 s_dak_total_prov_8	3	area_4 dau_total_4 dau_total_prov_4
Log Value Added	logValAdded	448	3	s_pop1990_4 s_dak_total_4 s_dak_total_prov_8	3	area_4 dau_total_4 dau_total_prov_4
Log Total Labor	ln_L	448	3	s_pop1990_4 s_dak_total_4 s_dak_total_prov_8	3	area_4 dau_total_4 dau_total_prov_4
Log Output per Worker	logOutputPerWorker	448	3	s_pop1990_4 s_dak_total_4 s_dak_total_prov_8	3	area_4 dau_total_4 dau_total_prov_4

Notes: Notes here.

Table A.19: ISLAND MARKET POTENTIAL AND FIRM-LEVEL MANUFACTURING OUTCOMES

	FELS	**Other District IV-Lasso**		**Own District IV-Lasso**		**Stats**	
	(1)	**(2)**	**KBP**	**(3)**	**KBP**	$\overline{Y}$	N
Log Output	-0.009	-0.412*	43.809	-0.241	28.554	24.338	322865.000
	(0.059)	(0.224)	(.)	(0.209)	(.)	(.)	(.)
		[0.19]		[1.00]			
Log Value Added	-0.075	-0.176	43.846	-0.221	28.567	9.085	322805.000
	(0.066)	(0.250)	(.)	(0.219)	(.)	(.)	(.)
		[0.33]		[1.00]			
Log Total Labor	-0.038	0.056	43.848	0.052	28.540	178.286	322982.000
	(0.023)	(0.081)	(.)	(0.078)	(.)	(.)	(.)
		[0.33]		[1.00]			
Log Output per Worker	0.031	-0.414**	43.823	2.402	28.560	0.085	322684.000
	(0.053)	(0.204)	(.)	(4.854)	(.)	(.)	(.)
		[0.19]		[1.00]			
Year FE	Yes	Yes		Yes			
District FE	Yes	Yes		Yes			
Firm FE	Yes	Yes		Yes			

Notes: We report the results of firm-level panel regressions of the dependent variable on island market potential. Each cell reports β from a separate regression, with the dependent variable listed in the row heading. All regressions include firm and year fixed effects (and implicitly also district fixed effects), with controls that include logs of current population and non-oil GDRP. Robust standard errors, clustered at the district level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.20: ROAD QUALITY AND FIRM-LEVEL MANUFACTURING: ROBUSTNESS (HAUSMAN IV)

		IV-Lasso							
	FELS	**Hausman**	**No Oil / Gas**	**No Mining**	**PLN**	**TV**	**Schools**	**Health Facil.**	**Only 1998+**
Panel A: IFLS Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Log Output	-0.031	-0.055	-0.115	-0.109	0.038	-0.074	-0.065	-0.073	-0.125
	(0.041)	(0.074)	(0.076)	(0.073)	(0.084)	(0.074)	(0.071)	(0.073)	(0.255)
N	278434	278434	225877	187627	277431	277431	277431	277431	154987
Kleibergen-Paap Wald Rank *F* Stat		22.23	26.64	43.29	20.58	22.43	26.95	24.84	15.19
Log Value Added	-0.055	0.017	-0.054	-0.109	0.100	0.018	0.020	0.003	-0.019
	(0.047)	(0.083)	(0.075)	(0.069)	(0.094)	(0.085)	(0.079)	(0.078)	(0.273)
N	278368	278368	225839	187596	277367	277367	277367	277367	154942
Kleibergen-Paap Wald Rank *F* Stat		22.21	26.56	43.20	20.56	22.40	26.94	24.78	15.20
Log Total Labor	-0.007	-0.016	-0.023	-0.046	-0.014	-0.016	-0.020	-0.010	-0.009
	(0.012)	(0.021)	(0.022)	(0.031)	(0.020)	(0.021)	(0.021)	(0.022)	(0.056)
N	278539	278539	225982	187720	277537	277537	277537	277537	155096
Kleibergen-Paap Wald Rank *F* Stat		22.20	26.57	43.15	20.55	22.39	26.93	24.78	15.20
Log Output per Worker	-0.023	-0.115	-0.122	-0.078	0.015	-0.123*	-0.055	-0.119*	-0.116
	(0.039)	(0.072)	(0.074)	(0.067)	(0.219)	(0.070)	(0.071)	(0.067)	(0.248)
N	278284	278284	225774	187553	277282	277282	277282	277282	154903
Kleibergen-Paap Wald Rank *F* Stat		22.23	26.65	43.27	20.57	22.42	26.94	24.84	15.18

Notes: We report the results of community-level panel regressions of the dependent variable on market potential (both in logs). Each cell reports β from a separate regression, with the dependent variable listed in the row heading. All regressions include community and year fixed effects, with controls including logs of current population and non-oil GDRP. Interpretation of results remains unchanged when dependent variables are expressed in levels. Dependent variable means are reported in levels. Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.21: ROAD QUALITY AND FIRM-LEVEL MANUFACTURING: ROBUSTNESS (B-TARGET IV)

		IV-Lasso							
	FELS	**bTarget**	**No Oil / Gas**	**No Mining**	**PLN**	**TV**	**Schools**	**Health Facil.**	**Only 1998+**
Panel A: IFLS Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Log Output	-0.031	-0.054	-0.080	-0.109	-0.079	-0.054	-0.021	-0.050	1.000
	(0.041)	(0.058)	(0.066)	(0.073)	(0.060)	(0.058)	(0.056)	(0.057)	(.)
N	278434	278434	225877	187627	277431	277431	277431	277431	282529
Kleibergen-Paap Wald Rank *F* Stat		37.24	42.88	43.29	31.43	37.70	39.39	41.52	
Log Value Added	-0.055	-0.025	-0.084	-0.109	-0.062	-0.025	0.011	-0.024	1.000
	(0.047)	(0.061)	(0.059)	(0.069)	(0.063)	(0.061)	(0.059)	(0.058)	(.)
N	278368	278368	225839	187596	277367	277367	277367	277367	282469
Kleibergen-Paap Wald Rank *F* Stat		37.22	42.73	43.20	31.38	37.67	39.38	41.41	
Log Total Labor	-0.007	-0.028	-0.031	-0.046	-0.024	-0.028	-0.028	-0.024	1.000
	(0.012)	(0.018)	(0.025)	(0.031)	(0.018)	(0.018)	(0.018)	(0.017)	(.)
N	278539	278539	225982	187720	277537	277537	277537	277537	282643
Kleibergen-Paap Wald Rank *F* Stat		37.20	42.76	43.15	31.38	37.65	39.37	41.41	
Log Output per Worker	-0.023	-0.030	-0.058	-0.078	-0.058	-0.031	0.016	-0.027	1.000
	(0.039)	(0.055)	(0.060)	(0.067)	(0.057)	(0.054)	(0.052)	(0.053)	(.)
N	278284	278284	225774	187553	277282	277282	277282	277282	282407
Kleibergen-Paap Wald Rank *F* Stat		37.22	42.87	43.27	31.41	37.68	39.36	41.50	

Notes: We report the results of community-level panel regressions of the dependent variable on market potential (both in logs). Each cell reports β from a separate regression, with the dependent variable listed in the row heading. All regressions include community and year fixed effects, with controls including logs of current population and non-oil GDRP. Interpretation of results remains unchanged when dependent variables are expressed in levels. Dependent variable means are reported in levels. Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.22: ROAD QUALITY AND FIRM-LEVEL MANUFACTURING OUTCOMES: HETEROGENEITY (HAUSMAN IV)

		IV-Lasso				
			GHSL 1990		Pop 1990	
	FELS	**Hausman**	**$<$ 20 pct.**	**$\geq$ 20 pct.**	**$<$ 1 mil.**	**$\geq$ 1 mil.**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Log Output	-0.031	-0.004	0.506***	-0.138*	0.351**	-0.084
	(0.041)	(0.081)	(0.146)	(0.083)	(0.166)	(0.083)
N	278434	278434	25468	252966	35654	242780
Kleibergen-Paap Wald Rank *F* Stat		34.67	115.42	26.39	10.32	28.23
Log Value Added	-0.055	0.045	0.526***	-0.088	0.373**	-0.033
	(0.047)	(0.089)	(0.162)	(0.096)	(0.179)	(0.092)
N	278368	278368	25446	252922	35638	242730
Kleibergen-Paap Wald Rank *F* Stat		34.64	115.57	26.36	10.30	28.20
Log Total Labor	-0.007	-0.018	0.022	-0.035	0.058	-0.030
	(0.012)	(0.022)	(0.037)	(0.025)	(0.050)	(0.025)
N	278539	278539	25471	253068	35660	242879
Kleibergen-Paap Wald Rank *F* Stat		34.65	115.42	26.38	10.31	28.21
Log Output per Worker	-0.023	-0.026	0.490***	-0.104	0.300**	-0.054
	(0.039)	(0.079)	(0.148)	(0.083)	(0.136)	(0.085)
N	278284	278284	25443	252841	35636	242648
Kleibergen-Paap Wald Rank *F* Stat		34.67	115.09	26.40	10.31	28.23

Notes: Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.23: ROAD QUALITY AND FIRM-LEVEL MANUFACTURING OUTCOMES: HETEROGENEITY (BTARGET IV)

		IV-Lasso				
			GHSL 1990		Pop 1990	
	FELS	**bTarget**	**< 20 pct.**	**≥ 20 pct.**	**< 1 mil.**	**≥ 1 mil.**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Log Output	-0.031	0.006	0.288**	-0.096	0.416***	-0.038
	(0.041)	(0.065)	(0.117)	(0.084)	(0.151)	(0.076)
N	278434	278434	25468	252966	35654	242780
Kleibergen-Paap Wald Rank F Stat		55.00	43.45	35.11	8.02	47.18
Log Value Added	-0.055	0.031	0.355**	-0.065	0.419**	-0.001
	(0.047)	(0.068)	(0.139)	(0.089)	(0.181)	(0.081)
N	278368	278368	25446	252922	35638	242730
Kleibergen-Paap Wald Rank F Stat		54.99	43.60	35.11	8.04	47.17
Log Total Labor	-0.007	-0.031*	0.021	-0.046**	0.092*	-0.041*
	(0.012)	(0.019)	(0.038)	(0.023)	(0.055)	(0.021)
N	278539	278539	25471	253068	35660	242879
Kleibergen-Paap Wald Rank F Stat		54.98	43.33	35.10	8.04	47.17
Log Output per Worker	-0.023	0.031	0.271**	-0.050	0.337***	0.003
	(0.039)	(0.062)	(0.116)	(0.081)	(0.115)	(0.076)
N	278284	278284	25443	252841	35636	242648
Kleibergen-Paap Wald Rank F Stat		54.98	43.41	35.12	8.01	47.17

Notes: Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.24: Selected IVs for Table 6: Effects of Road Quality on Consumption, Income, and Employment

			Hausman		Base Target	
	DepVar	# Total	# Sel.	Selected IVs	# Sel.	Selected IVs
Log HH Per-Capita Consumption Expenditures	lnpce_defl	672	2	lgdrp_conp s_pop1971_8 s_rd_slope8_N_2	6	pop1971_8 rd_slope2_P_7 rd_slope8_P_3 tot_length_P_9 dbh_total_prov_12 hhSa
Log Total Earnings	l2salary2f2	448	2	s_pop1971_8 s_rd_slope8_N_2	4	pop1971_8 d_majorports_8 rd_slope2_N_7 rd_slope2_P_7
Log Agricultural Earnings	l2agriflag2	448	1	s_pop1971_8	2	pop1971_8 rd_slope2_P_8
Log Total Hours Worked	whrs_mth_n	448	1	s_pop1971_8	3	pop1971_8 rd_slope2_N_7 rd_slope2_P_7
Any Employment (0 1)?	working	448	3	s_pop1971_8 s_rd_slope8_N_2 s_rd_slope8_N_6	6	pop1971_7 pop1971_8 d_majorports_8 rd_slope8_N_4 rd_slope2_P_7 rd_slope8_P
Manufacturing ... Any Employment (0 1)?	working_manu	448	1	s_pop1971_8	3	pop1971_8 rd_slope2_N_7 rd_slope2_P_7
Other, Formal ... Any Employment (0 1)?	working_otherFormal	448	1	s_pop1971_8	3	pop1971_8 rd_slope2_N_7 rd_slope2_P_7
Other, Informal ... Any Employment (0 1)?	working_otherInformal	448	1	s_pop1971_8	3	pop1971_8 rd_slope2_N_7 rd_slope2_P_7

Notes here.

Table A.25: Effects of Island Market Potential on Consumption, Income, and Employment

	FELS	Other District IV-Lasso		Own District IV-Lasso		Stats	
Panel A: Household-Level Outcomes	**(1)**	**(2)**	**KBP**	**(3)**	**KBP**	$\overline{Y}$	N
Log HH Per-Capita Consumption Expenditures	0.198***	0.323***	274.259	0.190**	75.590	11.096	24078.000
	(0.065)	(0.099)	(.)	(0.089)	(.)	(.)	(.)
		[0.01]		[0.11]			
Panel B: Individual-Level Outcomes							
Log Total Earnings	0.190***	0.181***	198.871	0.109	57.375	1.463	19012.000
	(0.045)	(0.053)	(.)	(0.072)	(.)	(.)	(.)
		[0.01]		[0.22]			
Log Agricultural Earnings	0.161*	0.439***	179.540	0.080	34.648	0.729	5506.000
	(0.096)	(0.136)	(.)	(0.201)	(.)	(.)	(.)
		[0.01]		[0.87]			
Log Total Hours Worked	-0.032	-0.034	193.900	-0.006	59.097	199.688	24628.000
	(0.052)	(0.073)	(.)	(0.103)	(.)	(.)	(.)
		[0.23]		[0.92]			
Any Employment (0 1)?	0.008	-0.012	241.737	0.012	63.226	0.697	39290.000
	(0.028)	(0.039)	(.)	(0.061)	(.)	(.)	(.)
		[0.23]		[0.92]			
Manufacturing ... Any Employment (0 1)?	0.134***	0.109**	194.113	0.219***	59.204	0.294	24632.000
	(0.032)	(0.043)	(.)	(0.056)	(.)	(.)	(.)
		[0.02]		[0.00]			
Other, Formal ... Any Employment (0 1)?	0.053*	0.081*	194.113	0.060	59.204	0.313	24632.000
	(0.031)	(0.045)	(.)	(0.071)	(.)	(.)	(.)
		[0.05]		[0.49]			
Other, Informal ... Any Employment (0 1)?	-0.161***	-0.159***	194.113	-0.277***	59.204	0.462	24632.000
	(0.030)	(0.041)	(.)	(0.062)	(.)	(.)	(.)
		[0.01]		[0.00]			
Individual FE	Yes	Yes	Yes	.	.	.	.
Year FE	Yes	Yes	Yes	.	.	.	.

Note: We report the results of individual-level panel regressions with individual and survey-wave fixed effects. Each cell reports estimates of β from a separate regression, with the dependent variable listed in the row heading. Controls include: district GDP, individual age, education, household size, and month of survey indicators. Total hours worked is defined only if the individual reported working. Earnings regressions also include hours worked (by sector) as a control. Robust standard errors in parentheses, clustered at the (initial) village level. */**/*** denotes significant at the 10% / 5% / 1% levels. Full results, including results restricted to the sample of non-moving individuals, can be found in Appendix Tables.

Table A.26: EFFECTS OF ROAD QUALITY ON CONSUMPTION, INCOME, AND EMPLOYMENT (NON-MOVERS)

	FELS	Other District IV-Lasso		Own District IV-Lasso		Stats	
Panel A: Household-Level Outcomes	**(1)**	**(2)**	**KBP**	**(3)**	**KBP**	$\overline{Y}$	N
Log HH Per-Capita Consumption Expenditures	0.120***	0.194**	140.013	0.016	41.490	11.050	20281.000
	(0.042)	(0.091)	(.)	(0.098)	(.)	(.)	(.)
		[0.06]		[1.00]			
Panel B: Individual-Level Outcomes							
Log Total Earnings	0.056*	0.161**	162.212	0.066	44.913	1.417	16368.000
	(0.033)	(0.064)	(.)	(0.062)	(.)	(.)	(.)
		[0.04]		[0.61]			
Log Agricultural Earnings	0.152**	0.434***	174.176	0.285*	42.753	0.716	5040.000
	(0.071)	(0.120)	(.)	(0.149)	(.)	(.)	(.)
		[0.01]		[0.16]			
Log Total Hours Worked	0.009	-0.090	173.153	-0.003	37.155	198.245	21312.000
	(0.039)	(0.079)	(.)	(0.091)	(.)	(.)	(.)
		[0.22]		[1.00]			
Any Employment (0 1)?	-0.035	-0.041	187.529	-0.018	32.213	0.704	33418.000
	(0.022)	(0.044)	(.)	(0.045)	(.)	(.)	(.)
		[0.27]		[1.00]			
Manufacturing ... Any Employment (0 1)?	0.085***	0.072	173.298	0.128**	37.208	0.289	21315.000
	(0.024)	(0.047)	(.)	(0.051)	(.)	(.)	(.)
		[0.13]		[0.08]			
Other, Formal ... Any Employment (0 1)?	0.006	0.033	173.298	0.006	37.208	0.309	21315.000
	(0.022)	(0.053)	(.)	(0.056)	(.)	(.)	(.)
		[0.30]		[1.00]			
Other, Informal ... Any Employment (0 1)?	-0.091***	-0.163***	173.298	-0.169***	37.208	0.472	21315.000
	(0.023)	(0.047)	(.)	(0.048)	(.)	(.)	(.)
		[0.01]		[0.02]			
Individual FE	Yes	Yes	Yes	.	.	.	.
Year FE	Yes	Yes	Yes	.	.	.	.

Note: We report the results of individual-level panel regressions with individual and survey-wave fixed effects. Each cell reports estimates of β from a separate regression, with the dependent variable listed in the row heading. Controls include: district GDP, individual age, education, household size, and month of survey indicators. Total hours worked is defined only if the individual reported working. Earnings regressions also include hours worked (by sector) as a control. Robust standard errors in parentheses, clustered at the (initial) village level. */**/*** denotes significant at the 10% / 5% / 1% levels. Full results, including results restricted to the sample of non-moving individuals, can be found in Appendix Tables.

Table A.27: Effects of Road Roughness on Consumption, Income, and Employment (Robustness, Hausman IV)

		IV-Lasso							
	FELS	**Hausman**	**No Oil / Gas**	**No Mining**	**PLN**	**TV**	**Schools**	**Health Facil.**	**Only 1998+**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Log HH Per-Capita Consumption Expenditures	0.132***	0.203**	0.205**	0.005	0.247***	0.196**	0.206**	0.210**	0.391***
	(0.042)	(0.087)	(0.090)	(0.102)	(0.086)	(0.092)	(0.086)	(0.093)	(0.097)
N	22036	22036	18243	15812	22036	22036	22036	22036	11726
Kleibergen-Paap Wald Rank *F* Stat		150.23	131.61	25.70	136.32	135.18	154.82	128.41	94.30
Log Total Earnings	0.070**	0.232***	0.222***	-0.031	0.243***	0.212***	0.229***	0.253***	0.321***
	(0.033)	(0.055)	(0.060)	(0.090)	(0.057)	(0.057)	(0.056)	(0.059)	(0.117)
N	17619	17619	14943	12732	17619	17619	17619	17619	7758
Kleibergen-Paap Wald Rank *F* Stat		120.18	112.67	22.69	110.66	112.44	123.36	111.80	42.31
Log Agricultural Earnings	0.160**	0.435***	0.360***	-0.104	0.438***	0.451***	0.431***	0.442***	1.066***
	(0.071)	(0.122)	(0.136)	(0.258)	(0.123)	(0.132)	(0.124)	(0.123)	(0.273)
N	5246	5246	4210	3450	5246	5246	5246	5246	2268
Kleibergen-Paap Wald Rank *F* Stat		175.48	135.28	27.87	161.02	146.82	161.63	172.21	58.41
Log Total Hours Worked	-0.004	-0.070	-0.085	-0.054	-0.057	-0.081	-0.076	-0.076	-0.008
	(0.038)	(0.077)	(0.079)	(0.107)	(0.076)	(0.077)	(0.076)	(0.080)	(0.110)
N	22931	22931	19129	16471	22931	22931	22931	22931	10502
Kleibergen-Paap Wald Rank *F* Stat		182.41	87.50	17.91	166.20	171.03	184.50	164.89	86.79
Any Employment (0 1)?	-0.025	-0.030	-0.027	0.019	-0.023	-0.037	-0.029	-0.050	0.093*
	(0.021)	(0.037)	(0.040)	(0.064)	(0.037)	(0.039)	(0.037)	(0.038)	(0.051)
N	36257	36257	30339	26072	36257	36257	36257	36257	17086
Kleibergen-Paap Wald Rank *F* Stat		101.36	132.62	20.93	92.29	88.41	105.08	89.16	67.13
Manufacturing ... Any Employment (0 1)?	0.084***	0.071	0.111**	0.123*	0.075	0.061	0.069	0.058	0.043
	(0.024)	(0.047)	(0.047)	(0.066)	(0.047)	(0.048)	(0.047)	(0.047)	(0.067)
N	22934	22934	19132	16474	22934	22934	22934	22934	10502
Kleibergen-Paap Wald Rank *F* Stat		182.55	87.53	17.96	166.42	171.16	184.64	164.91	86.79
Other, Formal ... Any Employment (0 1)?	0.014	0.040	0.054	-0.000	0.046	0.037	0.043	0.045	-0.002
	(0.021)	(0.052)	(0.047)	(0.059)	(0.052)	(0.054)	(0.052)	(0.054)	(0.076)
N	22934	22934	19132	16474	22934	22934	22934	22934	10502
Kleibergen-Paap Wald Rank *F* Stat		182.55	87.53	17.96	166.42	171.16	184.64	164.91	86.79
Other, Informal ... Any Employment (0 1)?	-0.095***	-0.154***	-0.177***	-0.139**	-0.165***	-0.149***	-0.156***	-0.159***	-0.068
	(0.022)	(0.046)	(0.046)	(0.058)	(0.046)	(0.048)	(0.046)	(0.048)	(0.064)
N	22934	22934	19132	16474	22934	22934	22934	22934	10502
Kleibergen-Paap Wald Rank *F* Stat		182.55	87.53	17.96	166.42	171.16	184.64	164.91	86.79

Notes: Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.28: Effects of Road Roughness on Consumption, Income, and Employment (Robustness, bTarget IV)

		IV-Lasso							
	FELS	**bTarget**	**No Oil / Gas**	**No Mining**	**PLN**	**TV**	**Schools**	**Health Facil.**	**Only 1998+**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Log HH Per-Capita Consumption Expenditures	0.132***	0.183**	0.141**	0.005	0.167**	0.169**	0.157*	0.210**	0.447***
	(0.042)	(0.081)	(0.071)	(0.102)	(0.081)	(0.083)	(0.080)	(0.085)	(0.085)
N	22036	22036	18243	15812	22036	22036	22036	22036	11726
Kleibergen-Paap Wald Rank F Stat		53.73	43.32	25.70	50.77	47.20	54.22	48.34	43.64
Log Total Earnings	0.070**	0.082	0.077	-0.031	0.063	0.068	0.073	-0.025	0.018
	(0.033)	(0.060)	(0.062)	(0.090)	(0.059)	(0.058)	(0.059)	(0.070)	(0.145)
N	17619	17619	14943	12732	17619	17619	17619	17619	7758
Kleibergen-Paap Wald Rank F Stat		34.72	27.38	22.69	38.49	39.83	35.22	28.64	26.86
Log Agricultural Earnings	0.160**	0.253*	0.360**	-0.104	0.265*	0.283*	0.236	0.270*	1.000
	(0.071)	(0.147)	(0.161)	(0.258)	(0.148)	(0.155)	(0.154)	(0.151)	(.)
N	5246	5246	4210	3450	5246	5246	5246	5246	5308
Kleibergen-Paap Wald Rank F Stat		43.96	38.33	27.87	41.05	42.70	38.72	38.36	
Log Total Hours Worked	-0.004	0.016	-0.041	-0.054	0.005	0.020	0.008	0.025	0.151
	(0.038)	(0.088)	(0.099)	(0.107)	(0.086)	(0.088)	(0.088)	(0.090)	(0.137)
N	22931	22931	19129	16471	22931	22931	22931	22931	10502
Kleibergen-Paap Wald Rank F Stat		47.28	31.74	17.91	50.59	51.87	47.42	37.56	25.09
Any Employment (0 1)?	-0.025	0.067	-0.023	0.019	0.068	0.071	0.066	0.017	0.257***
	(0.021)	(0.045)	(0.049)	(0.064)	(0.045)	(0.046)	(0.045)	(0.048)	(0.090)
N	36257	36257	30339	26072	36257	36257	36257	36257	17086
Kleibergen-Paap Wald Rank F Stat		29.35	26.13	20.93	28.63	33.67	30.51	24.80	18.08
Manufacturing ... Any Employment (0 1)?	0.084***	0.175***	0.120**	0.123*	0.169***	0.174***	0.177***	0.160***	0.140
	(0.024)	(0.053)	(0.056)	(0.066)	(0.052)	(0.053)	(0.052)	(0.054)	(0.105)
N	22934	22934	19132	16474	22934	22934	22934	22934	10502
Kleibergen-Paap Wald Rank F Stat		47.34	31.78	17.96	50.69	51.93	47.49	37.58	25.09
Other, Formal ... Any Employment (0 1)?	0.014	0.001	0.028	-0.000	-0.010	0.002	-0.002	-0.004	-0.051
	(0.021)	(0.056)	(0.062)	(0.059)	(0.055)	(0.056)	(0.057)	(0.057)	(0.093)
N	22934	22934	19132	16474	22934	22934	22934	22934	10502
Kleibergen-Paap Wald Rank F Stat		47.34	31.78	17.96	50.69	51.93	47.49	37.58	25.09
Other, Informal ... Any Employment (0 1)?	-0.095***	-0.181***	-0.153***	-0.139**	-0.164***	-0.181***	-0.177***	-0.180***	-0.029
	(0.022)	(0.050)	(0.056)	(0.058)	(0.048)	(0.050)	(0.051)	(0.053)	(0.085)
N	22934	22934	19132	16474	22934	22934	22934	22934	10502
Kleibergen-Paap Wald Rank F Stat		47.34	31.78	17.96	50.69	51.93	47.49	37.58	25.09

Notes: Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.29: Effects of Road Roughness on Consumption, Income, and Employment: Heterogeneity (Hausman IV)

		IV-Lasso				
			GHSL 1990		Pop 1990	
	FELS	Hausman	< 20 pct.	≥ 20 pct.	< 1 mil.	≥ 1 mil.
Panel A: Household-Level Outcomes	(1)	(2)	(3)	(4)	(5)	(6)
Log HH Per-Capita Consumption Expenditures	0.132***	0.203**	0.148	0.037	0.526***	0.080
	(0.042)	(0.087)	(0.133)	(0.115)	(0.151)	(0.105)
N	22036	22036	5845	16191	4669	17367
Kleibergen-Paap Wald Rank *F* Stat		150.23	107.29	84.80	29.21	122.06
Panel B: Individual-Level Outcomes	(1)	(2)	(3)	(4)	(5)	(6)
Log Total Earnings	0.070**	0.232***	0.157	0.220***	0.139*	0.227***
	(0.033)	(0.055)	(0.156)	(0.070)	(0.075)	(0.073)
N	17619	17619	3997	13622	3559	14060
Kleibergen-Paap Wald Rank *F* Stat		120.18	147.95	67.85	29.13	89.10
Log Agricultural Earnings	0.160**	0.435***	0.421*	0.465**	0.477	0.462***
	(0.071)	(0.122)	(0.248)	(0.192)	(0.402)	(0.143)
N	5246	5246	1998	3248	821	4425
Kleibergen-Paap Wald Rank *F* Stat		175.48	94.22	64.22	16.22	122.74
Log Total Hours Worked	-0.004	-0.070	-0.031	-0.137	0.167	-0.082
	(0.038)	(0.077)	(0.104)	(0.119)	(0.211)	(0.091)
N	22931	22931	6048	16883	4668	18263
Kleibergen-Paap Wald Rank *F* Stat		182.41	264.83	87.73	41.03	126.74
Any Employment (0 1)?	-0.025	-0.030	-0.083	-0.034	0.069	-0.050
	(0.021)	(0.037)	(0.052)	(0.051)	(0.066)	(0.051)
N	36257	36257	9196	27061	8105	28152
Kleibergen-Paap Wald Rank *F* Stat		101.36	106.12	57.33	21.90	90.79
Manufacturing ... Any Employment (0 1)?	0.084***	0.071	-0.023	0.011	0.036	0.075
	(0.024)	(0.047)	(0.067)	(0.065)	(0.093)	(0.058)
N	22934	22934	6048	16886	4668	18266
Kleibergen-Paap Wald Rank *F* Stat		182.55	264.83	87.83	41.03	126.83
Other, Formal ... Any Employment (0 1)?	0.014	0.040	0.016	0.058	0.081	0.034
	(0.021)	(0.052)	(0.064)	(0.082)	(0.108)	(0.062)
N	22934	22934	6048	16886	4668	18266
Kleibergen-Paap Wald Rank *F* Stat		182.55	264.83	87.83	41.03	126.83
Other, Informal ... Any Employment (0 1)?	-0.095***	-0.154***	-0.035	-0.148**	-0.159	-0.160***
	(0.022)	(0.046)	(0.065)	(0.070)	(0.104)	(0.056)
N	22934	22934	6048	16886	4668	18266
Kleibergen-Paap Wald Rank *F* Stat		182.55	264.83	87.83	41.03	126.83

Notes: Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.30: EFFECTS OF ROAD ROUGHNESS ON CONSUMPTION, INCOME, AND EMPLOYMENT: HETEROGENEITY (BTARGET IV)

		IV-Lasso				
			GHSL 1990		Pop 1990	
	FELS	**bTarget**	**$<$ 20 pct.**	**$\geq$ 20 pct.**	**$<$ 1 mil.**	**$\geq$ 1 mil.**
Panel A: Household-Level Outcomes	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Log HH Per-Capita Consumption Expenditures	0.132***	0.183**	0.065	-0.117	0.374***	0.033
	(0.042)	(0.081)	(0.146)	(0.109)	(0.145)	(0.103)
N	22036	22036	5845	16191	4669	17367
Kleibergen-Paap Wald Rank *F* Stat		53.73	22.73	28.67	23.14	46.93
Panel B: Individual-Level Outcomes	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Log Total Earnings	0.070**	0.082	-0.124	0.073	-0.055	0.018
	(0.033)	(0.060)	(0.132)	(0.081)	(0.103)	(0.070)
N	17619	17619	3997	13622	3559	14060
Kleibergen-Paap Wald Rank *F* Stat		34.72	30.36	22.85	16.47	28.89
Log Agricultural Earnings	0.160**	0.253*	-0.501*	0.550***	0.364	0.222
	(0.071)	(0.147)	(0.263)	(0.193)	(0.277)	(0.168)
N	5246	5246	1998	3248	821	4425
Kleibergen-Paap Wald Rank *F* Stat		43.96	22.33	29.12	9.28	34.48
Log Total Hours Worked	-0.004	0.016	-0.064	0.007	-0.281	0.087
	(0.038)	(0.088)	(0.134)	(0.129)	(0.229)	(0.102)
N	22931	22931	6048	16883	4668	18263
Kleibergen-Paap Wald Rank *F* Stat		47.28	32.05	25.57	11.23	38.79
Any Employment (0 1)?	-0.025	0.067	-0.007	0.018	0.108	-0.000
	(0.021)	(0.045)	(0.091)	(0.062)	(0.094)	(0.057)
N	36257	36257	9196	27061	8105	28152
Kleibergen-Paap Wald Rank *F* Stat		29.35	19.52	22.34	18.32	25.42
Manufacturing ... Any Employment (0 1)?	0.084***	0.175***	0.191***	0.104	0.084	0.147**
	(0.024)	(0.053)	(0.069)	(0.077)	(0.156)	(0.059)
N	22934	22934	6048	16886	4668	18266
Kleibergen-Paap Wald Rank *F* Stat		47.34	32.05	25.61	11.23	38.84
Other, Formal ... Any Employment (0 1)?	0.014	0.001	-0.046	0.042	0.027	0.003
	(0.021)	(0.056)	(0.061)	(0.087)	(0.148)	(0.062)
N	22934	22934	6048	16886	4668	18266
Kleibergen-Paap Wald Rank *F* Stat		47.34	32.05	25.61	11.23	38.84
Other, Informal ... Any Employment (0 1)?	-0.095***	-0.181***	-0.062	-0.197**	-0.210*	-0.166***
	(0.022)	(0.050)	(0.067)	(0.080)	(0.119)	(0.055)
N	22934	22934	6048	16886	4668	18266
Kleibergen-Paap Wald Rank *F* Stat		47.34	32.05	25.61	11.23	38.84

Notes: Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.31: Selected IVs for Table 7: Road Quality and District-Level Migration Outcomes

	DepVar	# Total	Hausman # Sel.	Hausman Selected IVs	Base Target # Sel.	Base Target Selected IVs
Percent Δ Population (2000-1990)	pctPopDelta	564	1	s_area_9	1	rd_slope5_N_11
Log Total Recent Migrants (Kabu)	log_migTotal	564	1	s_area_9	1	rd_slope5_N_11
Log Total Recent Migrants (Prov)	log_migTotal_prov	564	1	s_area_9	1	rd_slope5_N_11
Percent Δ Prov. Migrants (2000-1990)	pctMigDelta	564	1	s_slope6_11	1	rd_avgSlope_11

Notes here.

Table A.32: HEDONIC REGRESSIONS

	DV: Log Rent (1)	DV: Log Land Value (2)
Type of dwelling: Single Unit, Single Level	0.035	-0.093
	(0.040)	(0.113)
Type of dwelling: Single Unit, Multi Level	0.199***	-0.042
	(0.042)	(0.131)
Type of dwelling: Duplex	0.085*	0.010
	(0.046)	(0.138)
Type of dwelling: Multi Unit, Single Level	0.086*	0.217
	(0.048)	(0.183)
House is surrounded by human and animal waste	-0.016	-0.052
	(0.029)	(0.046)
House is surrounded by piles of trash	0.030	-0.044
	(0.021)	(0.042)
House is surrounded by stagnant water	-0.037*	-0.017
	(0.021)	(0.055)
There is a stable under / next to house	-0.037**	0.045
	(0.016)	(0.034)
House has sufficient ventilation	0.045**	0.035
	(0.018)	(0.036)
Owned house	0.109***	0.081
	(0.019)	(0.065)
House rented/contracted	-0.184***	0.143
	(0.035)	(0.241)
Yard is moderately sized	0.063***	0.173***
	(0.014)	(0.038)
Room number in the house	-0.087***	0.117***
	(0.004)	(0.009)
Ceramic floor	0.349***	0.409***
	(0.033)	(0.088)
Tiled floor	0.136***	0.334***
	(0.029)	(0.073)
Cement floor	0.038	0.197***
	(0.027)	(0.058)
Lumber floor	0.008	0.232***
	(0.039)	(0.080)
Bamboo floor	0.049	0.318**
	(0.081)	(0.147)
Masonry outer wall	0.235***	0.339***
	(0.028)	(0.061)
Lumber outer wall	0.075***	0.217***
	(0.028)	(0.056)
Concrete roof	0.304***	0.406
	(0.092)	(0.371)
Wooden roof	0.152***	0.126
	(0.057)	(0.151)
Metal roof	0.035	-0.058
	(0.039)	(0.088)
Tiled roof	0.099**	-0.058
	(0.042)	(0.086)
Asbestos roof	0.096*	-0.286
	(0.053)	(0.222)
Electricity in the house	0.097***	0.133**
	(0.028)	(0.064)
Piped water used for cooking	-0.011	-0.031
	(0.035)	(0.124)
Pump/Well water used for cooking	-0.094***	0.081
	(0.035)	(0.122)
Well/Spring/Rain water used for cooking	-0.176***	0.019
	(0.035)	(0.110)
River water used for cooking	-0.092*	-0.153
	(0.054)	(0.119)
Purchased water used for cooking	-0.294	-0.063
	(0.195)	(0.481)
Inside water source	0.063***	0.042
	(0.016)	(0.046)
Own toilet	0.106***	0.191***
	(0.017)	(0.040)
Drainage ditch (flowing)	0.069***	0.002
	(0.017)	(0.039)
Drainage ditch (stagnant)	0.038	0.024
	(0.025)	(0.045)
Trash collected by sanitation service	0.089***	0.192*
	(0.025)	(0.108)
N	28319	10692
Regression F-Stat	34.958	18.690
Adj. R^2	0.371	0.305
Adj. R^2 (Within)	0.055	0.075
Community × Wave FE	Yes	Yes

Note: *TO WRITE.* Robust standard errors in parentheses, clustered at the village level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.33: ISLAND MARKET POTENTIAL AND PRICES

	FELS	Other District IV-Lasso		Own District IV-Lasso		Stats	
	(1)	(2)	KBP	(3)	KBP	$\overline{Y}$	N
Log Factory Wage	0.127	0.321	50.506	0.182	95.069	3895.380	257.000
	(0.215)	(0.537)	(.)	(0.474)	(.)	(.)	(.)
		[0.61]		[1.00]			
Log Farm Wage	-0.100	-0.071	62.766	0.153	61.763	3908.402	373.000
	(0.154)	(0.206)	(.)	(0.201)	(.)	(.)	(.)
		[0.67]		[0.95]			
Log Food Price	-0.092	-0.010	233.798	-0.168	82.070	147.620	1045.000
	(0.080)	(0.112)	(.)	(0.147)	(.)	(.)	(.)
		[0.67]		[0.76]			
Log Perishables Price	-0.444***	-0.549***	233.798	-0.581***	82.070	76.639	1045.000
	(0.094)	(0.134)	(.)	(0.179)	(.)	(.)	(.)
		[0.02]		[0.13]			
Median Log Land Value	0.467	0.709*	219.720	-0.070	79.496	3970.338	886.000
	(0.337)	(0.415)	(.)	(0.601)	(.)	(.)	(.)
		[0.20]		[1.00]			
Median Log Rent	0.108	0.175	225.252	0.424**	82.751	3970.116	1057.000
	(0.118)	(0.140)	(.)	(0.199)	(.)	(.)	(.)
		[0.34]		[0.33]			
Log Land Value (Hedonic FE)	0.709**	1.370***	187.491	0.618	51.225	3904.616	683.000
	(0.274)	(0.348)	(.)	(0.417)	(.)	(.)	(.)
		[0.02]		[0.57]			
Log Rent (Hedonic FE)	0.197	0.408**	243.670	0.136	81.332	3967.263	1046.000
	(0.135)	(0.181)	(.)	(0.240)	(.)	(.)	(.)
		[0.13]		[0.96]			

We report the results of community-level panel regressions of the dependent variable on local road quality or market potential (both in logs). Each cell reports β from a separate regression, with the dependent variable listed in the row heading. Log(Farm Wage) is not available in 1993. Dependent variable means are reported in levels. Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.34: Effects of Road Roughness on Prices (Robustness, Hausman IV)

		IV-Lasso							
	FELS	**Hausman**	**No Oil / Gas**	**No Mining**	**PLN**	**TV**	**Schools**	**Health Facil.**	**Only 1998+**
Panel A: IFLS Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Log Factory Wage	-0.088	0.358	0.588	1.000	0.193	0.267	0.419	0.477	1.000
	(0.178)	(0.381)	(0.427)	(.)	(0.334)	(0.372)	(0.388)	(0.412)	(.)
N	226	226	186	226	226	226	226	226	226
Kleibergen-Paap Wald Rank *F* Stat		26.61	33.03		46.15	43.76	47.61	21.54	
Log Farm Wage	0.050	-0.072	-0.225	0.036	-0.101	-0.160	-0.099	1.000	1.000
	(0.115)	(0.172)	(0.184)	(0.359)	(0.175)	(0.190)	(0.165)	(.)	(.)
N	339	339	264	211	339	339	339	339	339
Kleibergen-Paap Wald Rank *F* Stat		128.02	107.44	29.47	118.75	99.24	124.22		
Log Food Price	-0.109	-0.015	0.178	0.203	-0.202	-0.220	-0.201	-0.268*	1.000
	(0.068)	(0.126)	(0.124)	(0.175)	(0.145)	(0.153)	(0.151)	(0.159)	(.)
N	914	914	771	684	907	907	907	907	914
Kleibergen-Paap Wald Rank *F* Stat		130.68	91.15	46.65	165.33	167.10	177.78	159.52	
Log Perishables Price	-0.314***	-0.635***	-0.570***	-0.233	-0.662***	-0.673***	-0.684***	-0.660***	1.000
	(0.079)	(0.139)	(0.139)	(0.162)	(0.198)	(0.203)	(0.202)	(0.212)	(.)
N	914	914	771	684	907	907	907	907	914
Kleibergen-Paap Wald Rank *F* Stat		130.68	91.15	46.65	165.33	167.10	177.78	159.52	
Median Log Land Value	0.557**	0.925**	0.944*	0.065	0.940**	0.873*	0.862**	0.942**	1.000
	(0.254)	(0.422)	(0.485)	(0.784)	(0.440)	(0.458)	(0.417)	(0.448)	(.)
N	778	778	650	563	771	771	771	771	778
Kleibergen-Paap Wald Rank *F* Stat		123.41	120.13	25.02	109.02	100.48	121.90	112.50	
Median Log Rent	0.116	-0.182	0.248	-0.240	-0.082	-0.139	-0.051	-0.081	1.000
	(0.091)	(0.218)	(0.157)	(0.201)	(0.185)	(0.196)	(0.179)	(0.195)	(.)
N	926	926	782	695	919	919	919	919	926
Kleibergen-Paap Wald Rank *F* Stat		190.65	95.31	40.15	124.82	124.19	132.50	119.92	
Log Land Value (Hedonic FE)	0.702***	1.447***	1.882***	0.072	1.547***	1.398***	1.453***	1.548***	1.000
	(0.218)	(0.390)	(0.407)	(0.816)	(0.422)	(0.419)	(0.390)	(0.404)	(.)
N	622	622	514	438	615	615	615	615	622
Kleibergen-Paap Wald Rank *F* Stat		160.81	94.34	34.24	136.20	139.63	154.56	148.65	
Log Rent (Hedonic FE)	0.081	-0.194	0.235	0.587**	-0.012	-0.082	0.022	-0.017	1.000
	(0.112)	(0.237)	(0.209)	(0.257)	(0.210)	(0.220)	(0.198)	(0.217)	(.)
N	914	914	770	683	908	908	908	908	914
Kleibergen-Paap Wald Rank *F* Stat		190.95	94.88	39.11	125.27	123.55	132.13	119.65	

Notes: Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.35: Effects of Road Roughness on Prices (Robustness, bTarget IV)

		IV-Lasso							
	FELS	**bTarget**	**No Oil / Gas**	**No Mining**	**PLN**	**TV**	**Schools**	**Health Facil.**	**Only 1998+**
Panel A: IFLS Travel Times	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Log Factory Wage	-0.088	0.201	0.469	1.000	-0.043	0.221	0.206	0.454	1.000
	(0.178)	(0.374)	(0.455)	(.)	(0.375)	(0.376)	(0.371)	(0.447)	(.)
N	226	226	186	226	226	226	226	226	226
Kleibergen-Paap Wald Rank *F* Stat		21.45	28.38		17.76	22.85	21.87	16.90	
Log Farm Wage	0.050	0.047	-0.033	0.036	-0.004	0.017	0.019	0.027	1.000
	(0.115)	(0.255)	(0.233)	(0.359)	(0.258)	(0.270)	(0.251)	(0.251)	(.)
N	339	339	264	211	339	339	339	339	339
Kleibergen-Paap Wald Rank *F* Stat		56.07	18.35	29.47	50.40	50.34	54.32	57.12	
Log Food Price	-0.109	0.050	0.022	0.203	0.115	0.066	0.106	-0.014	1.000
	(0.068)	(0.126)	(0.177)	(0.175)	(0.118)	(0.126)	(0.125)	(0.138)	(.)
N	914	914	771	684	907	907	907	907	914
Kleibergen-Paap Wald Rank *F* Stat		33.28	38.85	46.65	33.00	33.19	32.76	27.10	
Log Perishables Price	-0.314***	-0.372***	-0.142	-0.233	-0.395***	-0.401***	-0.382***	-0.344**	1.000
	(0.079)	(0.139)	(0.162)	(0.162)	(0.133)	(0.139)	(0.146)	(0.161)	(.)
N	914	914	771	684	907	907	907	907	914
Kleibergen-Paap Wald Rank *F* Stat		33.28	38.85	46.65	33.00	33.19	32.76	27.10	
Median Log Land Value	0.557**	-0.343	-0.404	0.065	-0.337	-0.337	-0.398	-0.462	1.000
	(0.254)	(0.556)	(0.681)	(0.784)	(0.560)	(0.571)	(0.578)	(0.601)	(.)
N	778	778	650	563	771	771	771	771	778
Kleibergen-Paap Wald Rank *F* Stat		58.78	29.00	25.02	56.22	60.88	57.47	45.90	
Median Log Rent	0.116	0.224	0.043	-0.240	0.290*	0.261	0.271	0.236	1.000
	(0.091)	(0.161)	(0.184)	(0.201)	(0.168)	(0.169)	(0.168)	(0.191)	(.)
N	926	926	782	695	919	919	919	919	926
Kleibergen-Paap Wald Rank *F* Stat		32.68	30.47	40.15	34.27	34.56	33.95	26.99	
Log Land Value (Hedonic FE)	0.702***	0.751*	0.835*	0.072	0.801*	0.712	0.816*	0.889**	1.000
	(0.218)	(0.408)	(0.461)	(0.816)	(0.429)	(0.435)	(0.424)	(0.433)	(.)
N	622	622	514	438	615	615	615	615	622
Kleibergen-Paap Wald Rank *F* Stat		42.86	37.27	34.24	36.78	42.41	40.61	36.52	
Log Rent (Hedonic FE)	0.081	0.233	0.160	0.587**	0.276	0.315	0.393*	0.179	1.000
	(0.112)	(0.204)	(0.244)	(0.257)	(0.202)	(0.202)	(0.207)	(0.214)	(.)
N	914	914	770	683	908	908	908	908	914
Kleibergen-Paap Wald Rank *F* Stat		37.42	34.97	39.11	38.07	37.70	36.02	30.08	

Notes: Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.36: EFFECTS OF ROAD ROUGHNESS ON PRICES (HETEROGENEITY, HAUSMAN IV)

		IV-Lasso				
			GHSL 1990		Pop 1990	
	FELS	Hausman	$<$ 20 pct.	$\geq$ 20 pct.	$<$ 1 mil.	$\geq$ 1 mil.
	(1)	(2)	(3)	(4)	(5)	(6)
Log Factory Wage	-0.088	0.358	0.901***	0.019	0.115	0.163
	(0.178)	(0.381)	(0.337)	(0.353)	(1.409)	(0.406)
N	226	226	29	197	41	185
Kleibergen-Paap Wald Rank *F* Stat		26.61	77.74	17.19	2.63	31.26
Log Farm Wage	0.050	-0.072	-0.059	-0.455	0.237	-0.016
	(0.115)	(0.172)	(0.353)	(0.367)	(0.441)	(0.226)
N	339	339	115	224	55	284
Kleibergen-Paap Wald Rank *F* Stat		128.02	65.66	25.62	11.50	89.72
Log Food Price	-0.109	-0.015	-0.734***	0.142	0.512**	-0.302*
	(0.068)	(0.126)	(0.215)	(0.154)	(0.224)	(0.164)
N	914	914	209	705	217	697
Kleibergen-Paap Wald Rank *F* Stat		130.68	83.88	77.92	16.53	124.69
Log Perishables Price	-0.314***	-0.635***	-0.311*	-0.636***	-0.574*	-0.681***
	(0.079)	(0.139)	(0.172)	(0.186)	(0.312)	(0.196)
N	914	914	209	705	217	697
Kleibergen-Paap Wald Rank *F* Stat		130.68	83.88	77.92	16.53	124.69
Median Log Land Value	0.557**	0.925**	0.226	1.222*	0.986	1.054**
	(0.254)	(0.422)	(0.482)	(0.624)	(1.996)	(0.488)
N	778	778	208	570	174	604
Kleibergen-Paap Wald Rank *F* Stat		123.41	67.58	65.11	15.70	98.08
Median Log Rent	0.116	-0.182	-0.947*	0.026	0.301	-0.353
	(0.091)	(0.218)	(0.563)	(0.233)	(0.300)	(0.265)
N	926	926	214	712	218	708
Kleibergen-Paap Wald Rank *F* Stat		190.65	169.02	90.02	22.91	162.04
Log Land Value (Hedonic FE)	0.702***	1.447***	0.026	2.384***	4.684***	1.287***
	(0.218)	(0.390)	(0.426)	(0.709)	(1.798)	(0.432)
N	622	622	192	430	126	496
Kleibergen-Paap Wald Rank *F* Stat		160.81	68.49	53.80	15.16	121.97
Log Rent (Hedonic FE)	0.081	-0.194	-0.684	-0.274	0.654*	-0.497*
	(0.112)	(0.237)	(0.476)	(0.315)	(0.366)	(0.289)
N	914	914	212	702	218	696
Kleibergen-Paap Wald Rank *F* Stat		190.95	167.14	90.54	22.91	162.34

Notes: Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.37: Effects of Road Roughness on Prices (Heterogeneity, bTarget IV)

		IV-Lasso				
			GHSL 1990		Pop 1990	
	FELS	**bTarget**	**$<$ 20 pct.**	**$\geq$ 20 pct.**	**$<$ 1 mil.**	**$\geq$ 1 mil.**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Log Factory Wage	-0.088	0.201	0.504	0.105	0.812	-0.038
	(0.178)	(0.374)	(0.546)	(0.383)	(0.748)	(0.484)
N	226	226	29	197	41	185
Kleibergen-Paap Wald Rank *F* Stat		21.45	9.41	17.09	12.52	17.74
Log Farm Wage	0.050	0.047	-0.328	-0.005	-0.100	0.043
	(0.115)	(0.255)	(0.395)	(0.442)	(0.549)	(0.305)
N	339	339	115	224	55	284
Kleibergen-Paap Wald Rank *F* Stat		56.07	31.85	20.93	5.38	43.42
Log Food Price	-0.109	0.050	-0.689***	-0.017	-0.579	-0.033
	(0.068)	(0.126)	(0.223)	(0.185)	(0.574)	(0.140)
N	914	914	209	705	217	697
Kleibergen-Paap Wald Rank *F* Stat		33.28	38.17	27.62	18.08	35.13
Log Perishables Price	-0.314***	-0.372***	-0.020	-0.599***	-0.850*	-0.368**
	(0.079)	(0.139)	(0.158)	(0.232)	(0.515)	(0.182)
N	914	914	209	705	217	697
Kleibergen-Paap Wald Rank *F* Stat		33.28	38.17	27.62	18.08	35.13
Median Log Land Value	0.557**	-0.343	-0.544	-0.663	-3.983	-0.130
	(0.254)	(0.556)	(0.632)	(0.867)	(3.063)	(0.592)
N	778	778	208	570	174	604
Kleibergen-Paap Wald Rank *F* Stat		58.78	38.53	28.31	3.83	48.67
Median Log Rent	0.116	0.224	-0.368	0.319	0.411	-0.168
	(0.091)	(0.161)	(0.285)	(0.204)	(0.351)	(0.206)
N	926	926	214	712	218	708
Kleibergen-Paap Wald Rank *F* Stat		32.68	35.42	24.39	18.05	34.95
Log Land Value (Hedonic FE)	0.702***	0.751*	-0.417	1.030	0.869	0.606
	(0.218)	(0.408)	(0.459)	(0.675)	(1.396)	(0.470)
N	622	622	192	430	126	496
Kleibergen-Paap Wald Rank *F* Stat		42.86	23.64	20.24	8.25	34.24
Log Rent (Hedonic FE)	0.081	0.233	-0.320	0.163	0.730**	-0.371
	(0.112)	(0.204)	(0.320)	(0.257)	(0.298)	(0.264)
N	914	914	212	702	218	696
Kleibergen-Paap Wald Rank *F* Stat		37.42	43.39	30.68	18.78	32.81

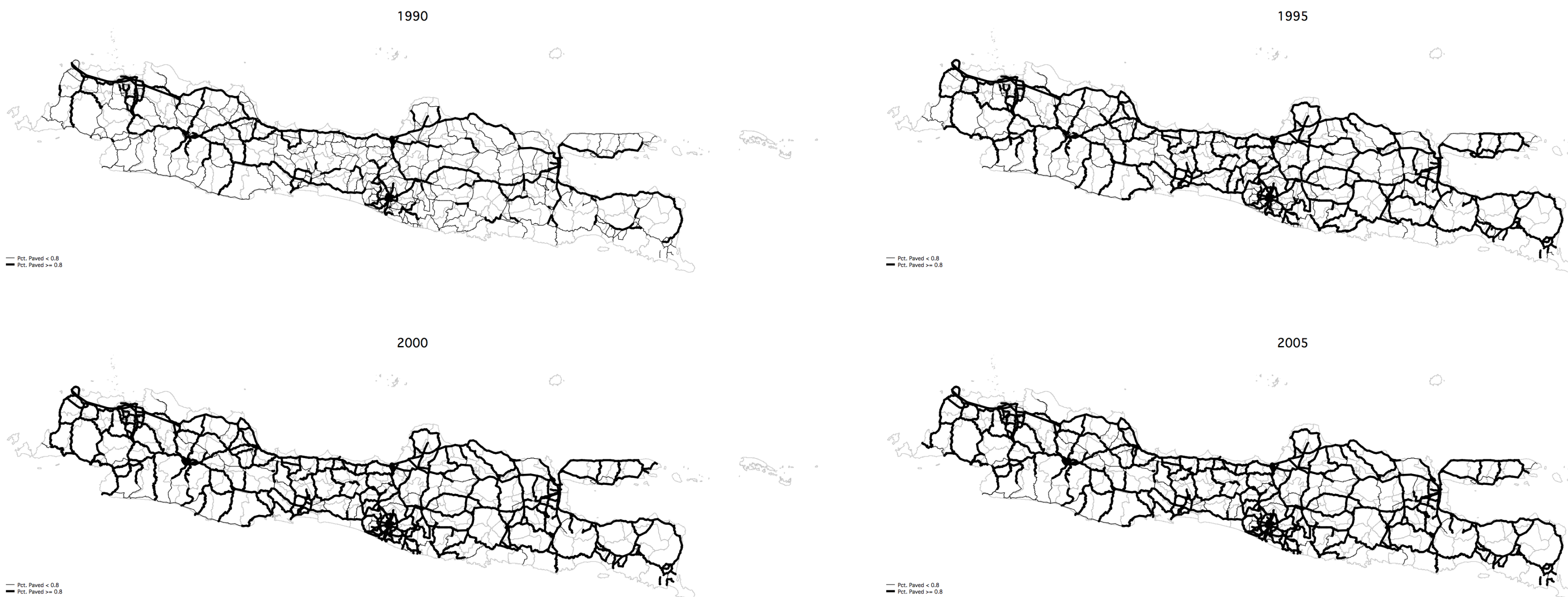
Notes: Robust standard errors, clustered at the community level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.38: SELECTED IVS FOR TABLE 8: PRICES AND ROAD QUALITY

			Hausman		Base Target	
	DepVar	# Total	# Sel.	Selected IVs	# Sel.	Selected IVs
Log Factory Wage	lwage_factd	672	2	s_pop1971_20 s_pop1990_20	2	pop1971_20 d_majorports_20
Log Farm Wage	lwage_farmd	672	1	s_pop1971_20	1	hhElectricity_prov_20
Log Food Price	lL_all1d	672	2	s_pop1971_20 s_rd_slope8_N_14	7	pop1971_19 pop1971_20 slope2_19 d_majorports_20 rd_slope2_P_19 dbh_total_prov_24 dak_total_prov_22
Log Tradables Price	lL_trade1d	672	2	s_pop1971_20 s_rd_slope8_N_14	7	pop1971_19 pop1971_20 slope2_19 d_majorports_20 rd_slope2_P_19 dbh_total_prov_24 dak_total_prov_22
Log Perishables Price	lL_perish1d	672	2	s_pop1971_20 s_rd_slope8_N_14	7	pop1971_19 pop1971_20 slope2_19 d_majorports_20 rd_slope2_P_19 dbh_total_prov_24 dak_total_prov_22
Log Factory Wage	lwage_factd	672	1	s_rd_slope1_N_20	1	dbh_total_prov_24
Log Farm Wage	lwage_farmd	672	1	s_pop1971_20	1	hhElectricity_prov_20
Log Food Price	lL_all1d	672	5	s_wat_2000_22 s_slope2_20 s_rd_slope1_N_20 s_rd_slope8_N_13 s_rd_slope8_N_14	5	pop1971_20 rd_slope2_N_19 rd_slope2_N_20 rd_slope2_N_22 dbh_total_prov_24
Log Tradables Price	lL_trade1d	672	5	s_wat_2000_22 s_slope2_20 s_rd_slope1_N_20 s_rd_slope8_N_13 s_rd_slope8_N_14	5	pop1971_20 rd_slope2_N_19 rd_slope2_N_20 rd_slope2_N_22 dbh_total_prov_24
Log Perishables Price	lL_perish1d	672	5	s_wat_2000_22 s_slope2_20 s_rd_slope1_N_20 s_rd_slope8_N_13 s_rd_slope8_N_14	5	pop1971_20 rd_slope2_N_19 rd_slope2_N_20 rd_slope2_N_22 dbh_total_prov_24

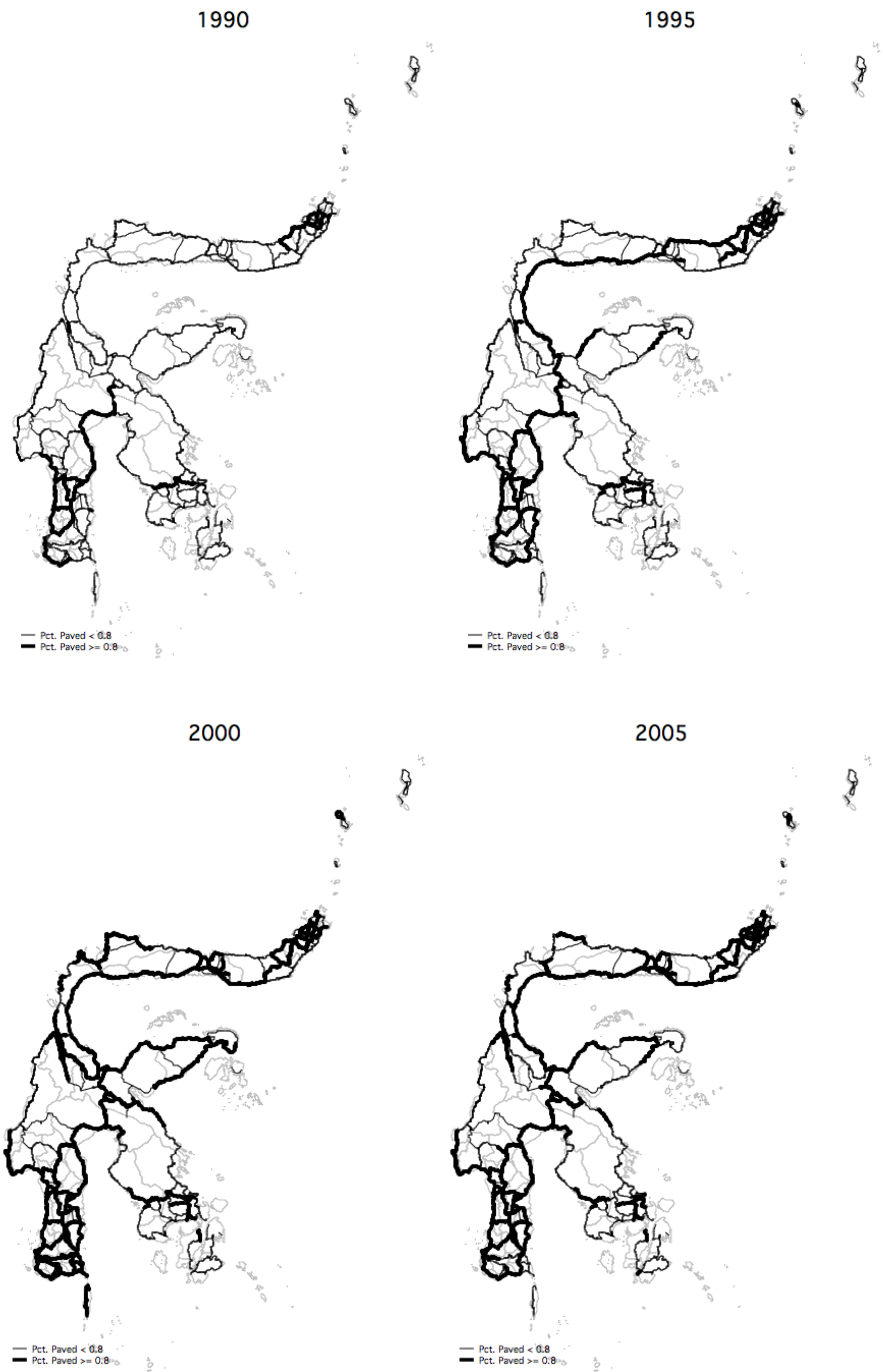
Notes here.

Figure A.1: Road Roughness in Java



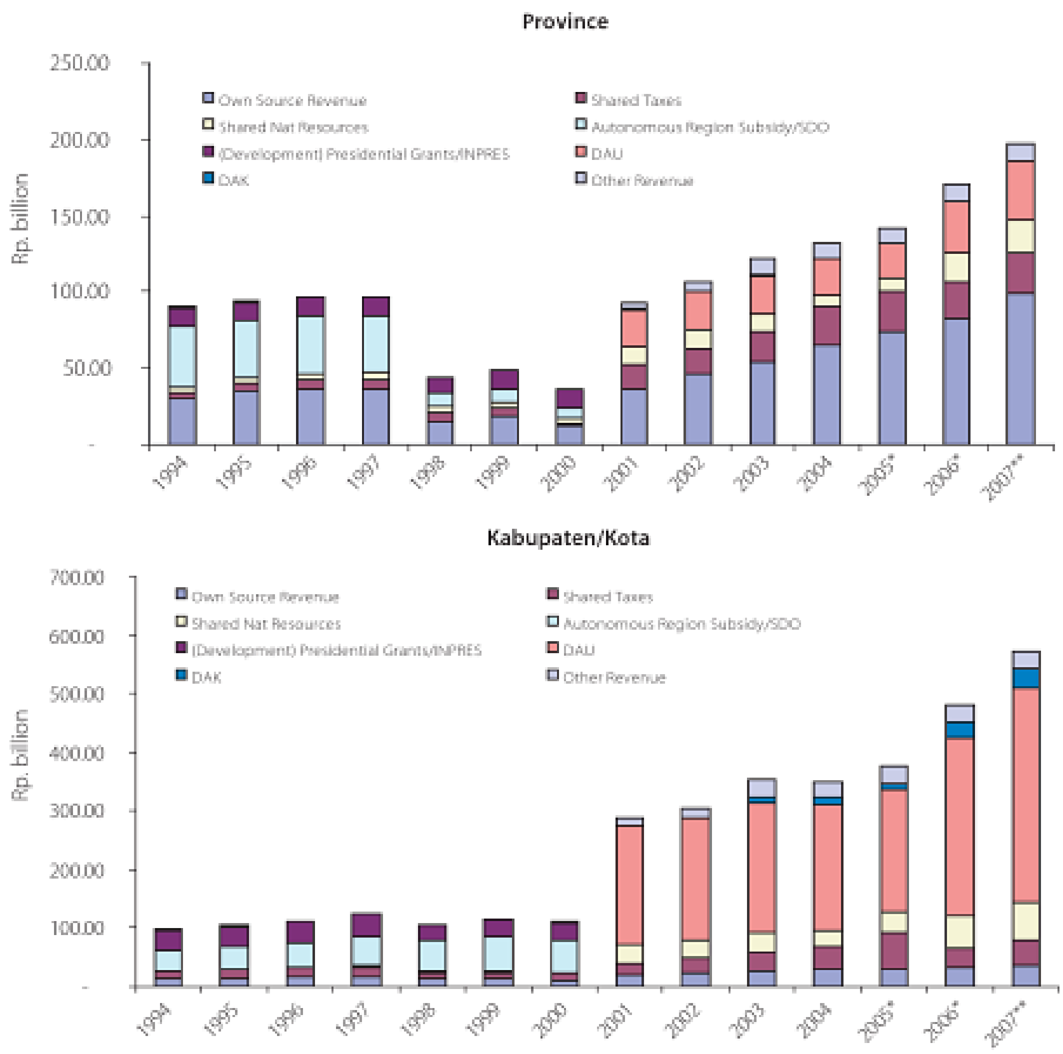
Note: Authors' calculations.

Figure A.2: Road Roughness in Sulawesi



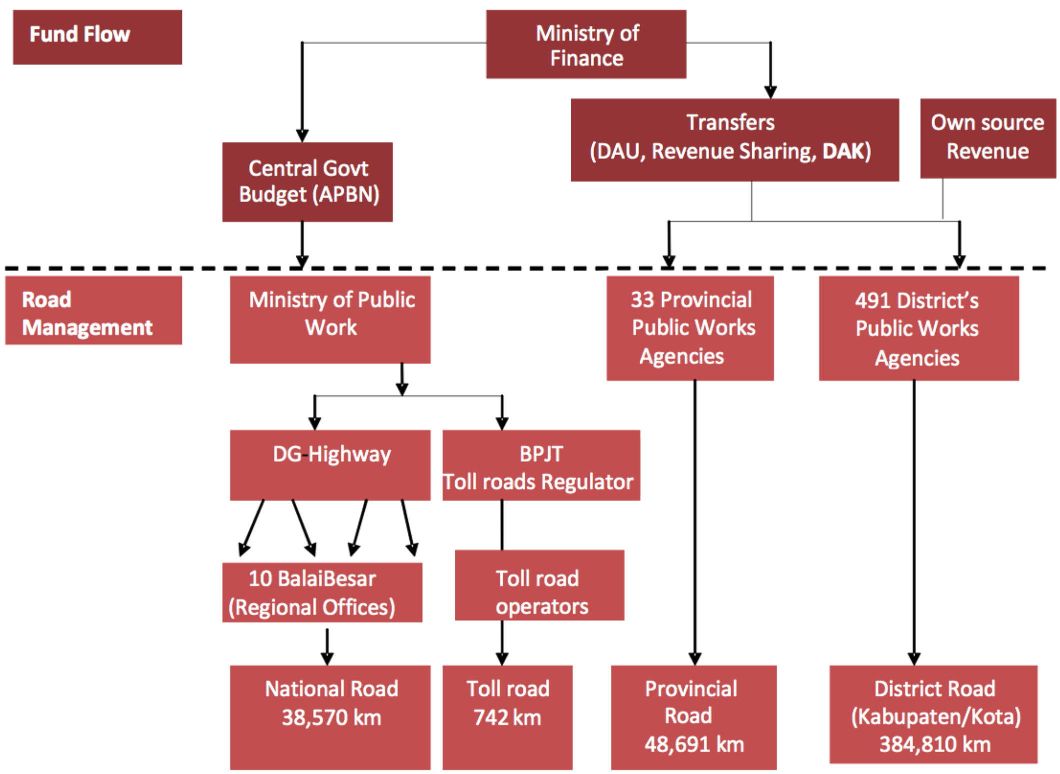
Note: Authors' calculations.

Figure A.3: Sub-national Revenue over Time



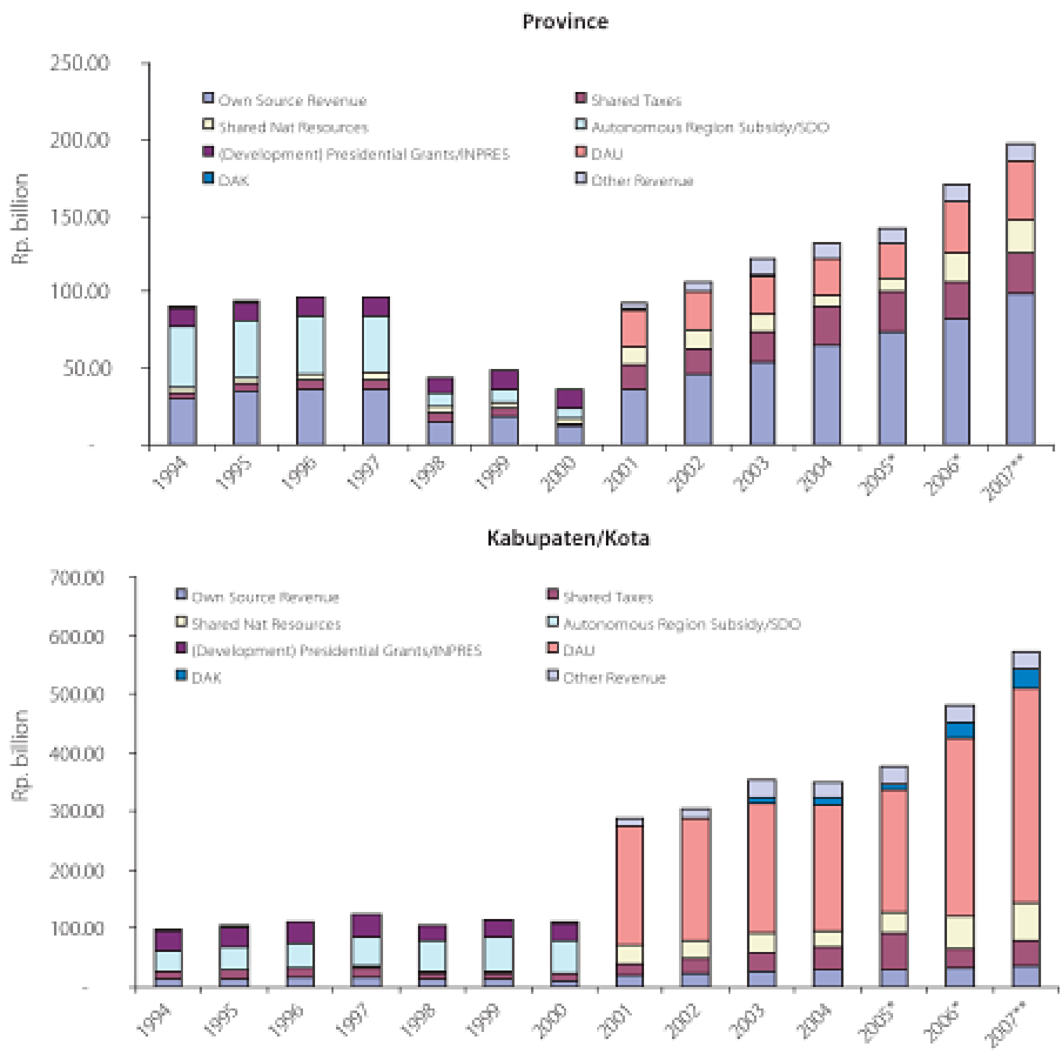
Note: World Bank staff calculations. (World Bank, 2008).

Figure A.4: Institutional Arrangements for Indonesia's Road Sector



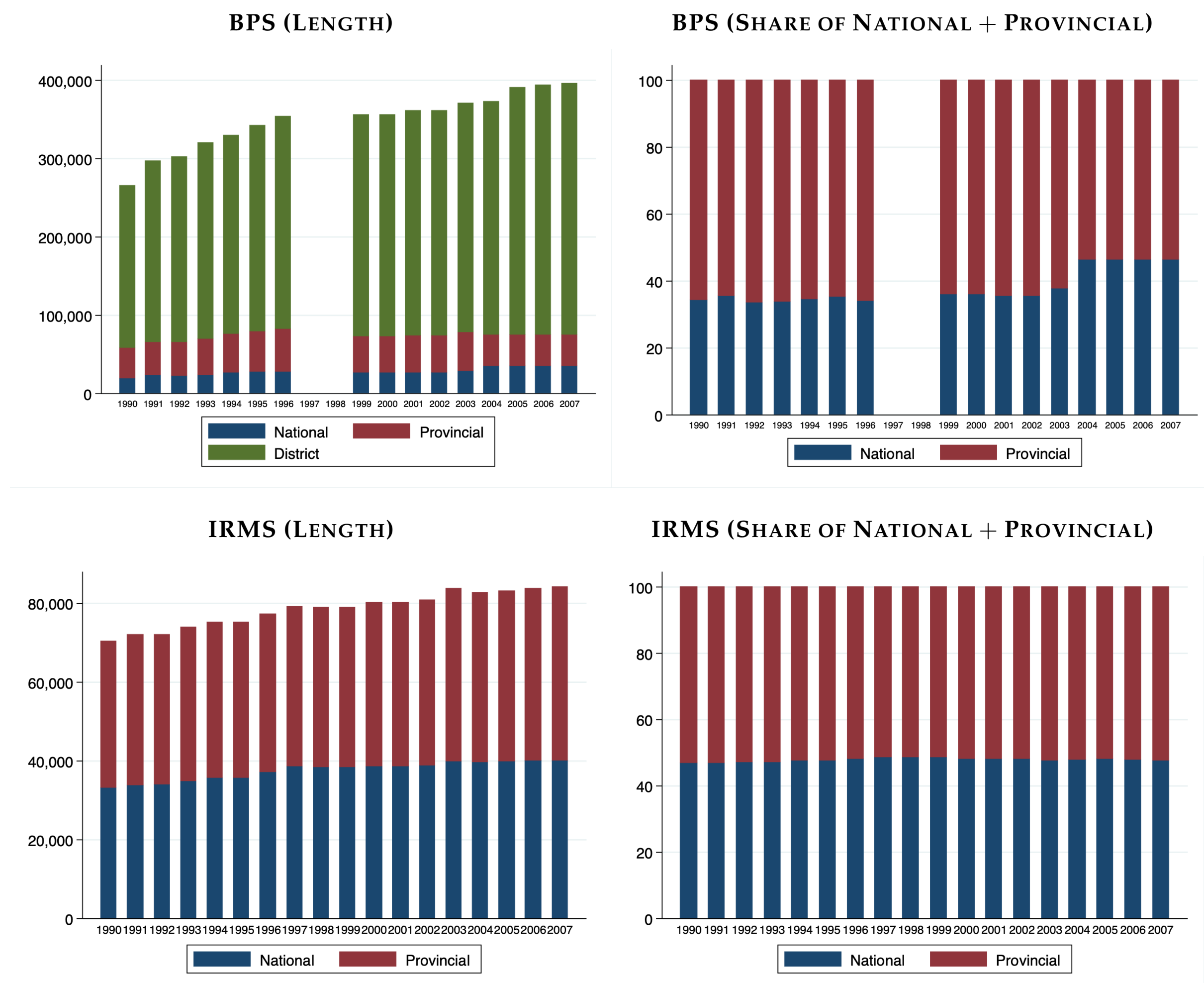
Source: (World Bank, 2012).

Figure A.5: Sub-national revenue over time



Note: World Bank staff calculation. World Bank (2008).

Figure A.6: Road Length by Administration Status



Note: Panels A and B are from various editions of BPS publications on National Transportation Statistics. These panels cover all roads in Indonesia. Panels C and D are from IRMS data, which only cover national and provincial roads on Java, Sumatra, and Sulawesi.

B Data Appendix

B.1 Road Quality Data

Data on the quality of Indonesia's highway networks were produced by DPU as part of Indonesia's Integrated Road Management System (IRMS). This appendix section begins by providing some background on road management in Indonesia, describing the road classification system and discussing IRMS coverage. It then discusses the measures of road quality that are collected in IRMS and how they are measured. I then discuss how the road network data were created.

B.1.1 Background on Road Management

Indonesia's national road network is currently managed and maintained by the Department of Public Works (*Departemen Pekerjaan Umum*, DPU), specifically by the Directorate General of Highways (*Direktorat Jenderal Bina Marga*). According to Law No. 38, 2004, roads are classified into four different types of roads, primarily based on their function for users. Arterial roads (*jalan arteri*) serve as the major transportation linkages between urban areas, and are characterized by longer distances, higher speeds, and limited access. Speeds are meant to be a minimum of 60 km/h, and width should be at least 11 meters to accommodate larger traffic volumes. Collector roads (*jalan kolektor*) serve "collector or distributor transportation" and are characterized by medium distance travel with medium speeds. Collector roads are subdivided into primary collector roads (*jalan kolektor primer*), which should have a minimum speed of 20 km/h and width of 9 meters, and secondary collector roads, which should have a minimum speed of 20 km/h and width of 9 meters. Local roads (*jalan lokal*) and Neighborhood Roads (*jalan lingkungan*) serve local areas at lower speeds, and are characterized by unlimited access.

Roads can also be classified by their management authority, or "status" (*wewenang penyelenggaraan*). Generally, arterial and primary collector roads are managed by the national government (specifically by DPU). Secondary and tertiary collector roads are managed by provincial governments, while local and neighborhood roads are managed by the kabupaten, kecamatan, and desa governments. Table B.1 describes the road classification system, minimum speed and width guidelines, and management authorities.

Table B.2 depicts the coverage of the IRMS dataset by road function and managing authority, as measured by counts of the number of kilometer-post observations that appear in the entire dataset. Most of the observations, and indeed most of the road network, is made up by collector roads (K1-K3), though the category with the next largest coverage is the arterial roads. Local and neighborhood roads are not very well surveyed in this dataset. Although the network of village and kabupaten roads is doubtless extremely dense, I cannot use this dataset to say very much about it. But since the data do cover arterial and collector roads, the major roads connecting regions and cities in Indonesia, this dataset seems particularly well suited for evaluating models of economic geography and regional trade.

B.1.2 Measures of Road Quality

There are a number of different devices that transport engineers have developed to collect measurements of road quality, and there are several different measures of road quality. The most widely used measure of road roughness, and the measure used in this study, is the international roughness index (IRI), developed by the World Bank in the 1980s. IRI is constructed as a filtered ratio of a standard vehicle's accumulated suspension motion (in meters), divided by the distance travelled by the vehicle during measurement (in kilometers). Expressed in units of slope (m/km), IRI is a characteristic of a vehicle's longitudinal profile. Importantly, since it is a measure of a physical quantity, IRI is standardized, as opposed to other subjective measures of ride quality. Figure **??** shows the relationship between different ranges of IRI and surface type; generally, larger roughness levels correspond to worse surfaces, but the mapping is not one-to-one.

Bennett et al. (2007) distinguish between several different types of devices for measuring road roughness and provide a good overview of their relative strengths and weaknesses. Over the course of its existence, Indonesia's IRMS has largely made use of two different types of measuring devices.[26] Before 1999, roads were surveyed using

[26] I am very grateful for the extensive discussions I've had with Glen Stringer about IRMS; this section of the appendix benefits highly from our conversations.

devices like the ROMDAS, which estimate IRI indirectly. The ROMDAS machine is a calibrated bump integrator, which must first be calibrated and estimates IRI from correlation equations. It is very useful for measuring roughness on bumpy roads and can record high levels of IRI, but the device must be calibrated manually, and measurement error can occur if the device is miscalibrated.

The ROMDAS device is also portable, meaning that it can be used inside different vehicles (each of which would require unique calibrations). The portability contrasts with devices like the high-speed laser profilometer, which is essentially a separate vehicle reserved entirely for the purposes of collecting road quality data. The device uses lasers and optical techniques to scan the road as it is traversed and create measures of surface profiles. These instruments are very accurate, but are much more expensive. Moreover, they might become mis-calibrated on extremely rough roads. Indonesia started using the high speed laser profilometer for collecting its road quality data in 1999, licensing vehicles from the Australian government.

Road width and surface type are more straightforward variables to measure, involving visual inspection and simple measurement. I categorize a kilometer-post interval as being unpaved if it is either an earth, gravel, or sand road, or if it was given a granular base (crushed stone) treatment, a first step in the process of paving.

B.1.3 Creation of Road Network Data

Using GIS shapefiles of the road network provided to me by DPU, I have georeferenced the kilometer post observations of road quality, in order to capture the evolution of Indonesia's transportation network over space and time. This proved to be a challenging exercise, because the identifiers for each road-link-interval observation were not consistent over time, and because the identifiers in the shapefile and in the linearly referenced dataset were often different, even though both did refer to exactly the same link.

Once the IRMS interval data was successfully merged to the regional network shapefiles, I converted the GIS database of road links into a weighted graph of arcs and nodes, as commonly used in the transportation literature. Nodes represent locations (such as ports, cities, or the centroids of kabupatens, my unit of analysis), arcs represent the possibility of traveling between two nodes, and weights represent the cost of moving goods along a given arc. Weights were constructed according to the IRMS data on road quality, and for simplicity, the cost of moving along each road was assumed to be the same, no matter which way you were traveling.[27]

For computational reasons, I have used a simplified representation of Indonesia's road network, where the number of nodes and links was small enough for network algorithms to operate on it using a desktop computer.[28] Table B.3 depicts the number of network arcs, the total distance of the network, and merge statistics for the kilometer-post observations. Merge statistics are pretty good for arterial and collector roads, but the quality of merges falls substantially for local and neighborhood roads, due most likely to poor shapefile coverage for that type of road network.

The interval observations were not matched directly to their exact locations in the network, because I had no knowledge of the exact location of the kilometer posts. To deal with this, I first aggregated the kilometer-post interval observations to the road-link level by constructing distance-weighted averages of the road quality variables. Each network arc-year observation was then assigned the value of this average road quality variable that corresponds to its road link.[29]

[27] Another tedious issue involved the construction of junction points where the road links intersected. The shapefiles were originally stored as MapInfo files, an older shapefile format that required conversion for use with Arcview, and in this conversion, information on where the roads crossed was lost, requiring painstaking editing. The shapefiles were also not designed to be used in any network analysis, so much care had to be taken to make them usable.

[28] The road lines were straightened using the "Generalize" command from ET Geotools, which employs the Ramer–Douglas–Peucker algorithm for reducing the number of points that represents a line.

[29] In some cases, when a network arc had no data for a particular year, I assigned the network arc the average value of road quality for arcs with the same function. This was done because constructing the transport cost variables involved a search over the entire network, and if certain network arcs were coded as missing, this could distort the search substantially. Overall, imputation amounted to no more than 5 percent of network arc observations in any given year.

B.1.4 Roughness, Speed, and Ride Quality

One effect that rough roads have on vehicles is that they require the driver to travel at lower speeds. When faced with potholes, ragged pavement, or poor surfaces, drivers slow down, and this reduction in speed increases travel time and hence the cost of travel. Of course, there is not a one-to-one relationship between road roughness and speed, because drivers choose the speed at which they travel, and different preferences for smoothness of the ride or the desired arrival time might induce different choices of speed.

Yu et al. (2006) explore the relationship between *jolt*, or the "jerk" experienced by road users, and subjective measures of ride quality and road roughness at different speeds.[30] Using survey data in which users were asked to rate the quality of particular rides, the authors find that people experience greater discomfort while traveling at higher speeds on rough roads, but lowering speed on rough roads can reduce discomfort. The authors provide a mapping between subjective measures of ride quality and roughness at different speeds, and this mapping can be used to infer the maximum speed that one can travel in order to achieve a ride of a certain quality, given pavement roughness. Table B.4 reproduces this mapping. Because travel times were unreasonably long for high quality rides given Indonesia's rough roads, and because the subjective quality measures were chosen by Western drivers, I have focused on the poor ride quality speed thresholds in my empirical work.

Given the maximum speed that one can travel on roads of different roughness levels, it is straightforward to calculate travel times for each network arc, the primary measure of transport costs used in this study. Note that the travel times on road sections were computed using the detailed kilometer-post interval roughness data. These were then aggregated to the network arcs using distance-weighted averages.

B.2 Administrative Boundaries

Administrative boundary shapefiles were constructed by BPS for use during the 2000 Household Census. These shapefiles contain the polygon boundaries of all provinces, kabupatens, kecamatans, and desas for the entire extent of the Indonesian archipelago. However, after the fall of Suharto and a massive decentralization program, many new kabupatens were created, splitting existing kabupatens into new ones. For instance, in 1990 there were 290 kabupatens and kotas, but by 2003, there were 416 kabupatens and kotas. The fact that administrative boundaries are not fixed over time create difficulties for the analysis.

Because of the need for a geographic unit of analysis that was consistently defined over time, I used kabupaten borders as they were defined in 1990. BPS provided the administrative boundary shapefile for 2000, as well as a correspondence table between kabupaten codes in 2000 and kabupaten codes from 1990 to the present. This information was processed using ArcView to create the 1990 shapefiles that form the basis of the analysis. Throughout the paper, all survey data were appropriately merged back to the 1990 kabupaten definitions.

[30] *Jolt* is officially defined as the vector that specifies the time-derivative of acceleration; in other words, the third derivative of the vertical displacement of vehicle to time t.

Table B.1: Indonesia's Road Classification System

Function	Code	Minimum Speed	Minimum Width	Management Authority
Arterial	A	60 km/h	11 m	National
Collector-1	K1	40 km/h	9 m	National
Collector-2	K2	20 km/h	9 m	Provincial
Collector-3	K3	20 km/h	9 m	Provincial
Local	L	20 km/h	7.5 m	Kabupaten & Desa
Neighborhood	Z	15 km/h	6.5 m	Kabupaten & Desa

Source: Departemen Pekerjaan Umum, 2008

Table B.2: Road Function and Managing Authority, Kilometer-Post Observations, 1990-2007

	Road Function			Managing Authority		
	Code	Number of Obs.	Share of Total	Code	Number of Obs.	Share of Total
Java	**A**	52,917	0.17	**N**	93,808	0.30
	K1	40,889	0.13	**P**	132,649	0.42
	K2	121,386	0.39	**K**	15,862	0.05
	K3	10,714	0.03	**S**	72,068	0.23
	L	15,862	0.05			
	Z	72,619	0.23			
	Total	314,387	1.00	**Total**	314,387	1.00
Sumatra	**A**	103,160	0.20	**N**	202,915	0.39
	K1	99,782	0.19	**P**	263,409	0.50
	K2	235,750	0.45	**K**	11,391	0.02
	K3	27,632	0.05	**S**	45,680	0.09
	L	11,391	0.02			
	Z	45,680	0.09			
	Total	523,395	1.00	**Total**	523,395	1.00
Sulawesi	**A**	54,496	0.21	**N**	143,147	0.54
	K1	87,728	0.33	**P**	72,198	0.27
	K2	71,234	0.27	**K**	18,232	0.07
	K3	1,887	0.01	**S**	29,371	0.11
	L	18,232	0.07			
	Z	29,371	0.11			
	Total	262,948	1.00	**Total**	262,948	1.00

Source: IRMS and author's calculations. Data come from kilometer-post observations. Standard deviations in parentheses.

Table B.3: Number of Network Arcs, Distances, and Merge Statistics (by road function)

		Road Function						
		A	K1	K2	K3	L	Z	Miss
Java	# of Arcs	1168	889	2618	309	315	37	.
	# of Road IDs	220	129	354	43	72	6	.
	Total Distance	2944.91	1970.65	5832.59	750.39	663.44	92.16	.
	Link-Years Merged	16538	13685	38719	3876	4689	14572	3015
	Link-Years Unmerged	1838	735	1842	45	971	21772	157
	% Merged	0.90	0.95	0.95	0.99	0.83	0.40	0.95
	Arc-Years Merged	20,844	16002	46350	5562	5670	666	.
	Arc-Years Unmerged	180	0	774	0	0	0	.
	% Merged	0.99	1.00	0.98	1.00	1.00	1.00	.
Sumatra	# of Arcs	1485	1205	2975	453	277	22	41
	# of Road IDs	207	165	412	87	66	6	13
	Total Distance	4964.69	4469.43	11551.28	1492.97	571.67	56.44	147.56
	Link-Years Merged	24755	20035	49171	6808	2603	8730	1406
	Link-Years Unmerged	718	373	537	52	394	9722	12
	% Merged	0.97	0.98	0.99	0.99	0.87	0.47	0.99
	Arc-Years Merged	26730	21690	51876	7830	4986	396	0
	Arc-Years Unmerged	0	0	1674	324	0	0	738
	% Merged	1.00	1.00	0.97	0.96	1.00	1.00	0.00
Sulawesi	# of Arcs	1624	2319	2051	15	391	.	45
	# of Road IDs	113	116	150	4	44	.	1
	Total Distance	2836.96	3805.92	4369.33	28.35	732.96	.	70.34
	Link-Years Merged	24006	24006	34711	30911	551	5670	5674
	Link-Years Unmerged	25	356	410	339	9	118	4755
	% Merged	1.00	0.99	0.99	0.99	0.98	0.98	0.54
	Arc-Years Merged	25794	35694	33660	270	7038	.	0
	Arc-Years Unmerged	3438	6048	3258	0	0	.	810
	% Merged	0.88	0.86	0.91	1.00	1.00		0.00

Source: IRMS and author's calculations. Missing function information is attributable to poorly coded shapefiles. Arc-Years could be unmerged potentially because there were no surveys done on that particular link; statistics are computed assuming a balanced panel. Road IDs are defined in the shapefile, while Link IDs are defined from the IRMS data.

Table B.4: Roughness and Ride-Quality Speed Limits

MAX SPEED	GOOD	FAIR	MEDIOCRE	POOR
120 KM/H	$IRI \in [0.00, 1.49]$	$IRI \in [0.00, 1.89]$	$IRI \in [0.00, 2.70]$	$IRI \in [0.00, 3.24]$
100 KM/H	$IRI \in [1.49, 1.79]$	$IRI \in [1.89, 2.27]$	$IRI \in [2.70, 3.24]$	$IRI \in [3.24, 4.05]$
80 KM/H	$IRI \in [1.79, 2.24]$	$IRI \in [2.27, 2.84]$	$IRI \in [3.24, 4.05]$	$IRI \in [4.05, 4.63]$
70 KM/H	$IRI \in [2.24, 2.57]$	$IRI \in [2.84, 3.25]$	$IRI \in [4.05, 4.63]$	$IRI \in [4.63, 5.40]$
60 KM/H	$IRI \in [2.57, 2.99]$	$IRI \in [3.25, 3.79]$	$IRI \in [4.63, 5.40]$	$IRI \in [5.40, 6.25]$
50 KM/H	$IRI \in [2.99, 3.59]$	$IRI \in [3.79, 4.54]$	$IRI \in [5.40, 6.25]$	$IRI \in [6.25, 8.08]$
40 KM/H	$IRI \in [3.59, 4.49]$	$IRI \in [4.54, 5.69]$	$IRI \in [6.25, 8.08]$	$IRI \in [8.08, 10.80]$
30 KM/H	$IRI \in [4.49, 5.99]$	$IRI \in [5.69, 7.59]$	$IRI \in [8.08, 10.80]$	$IRI \in [10.80, 16.16]$
20 KM/H	$IRI \in [5.99, 8.99]$	$IRI \in [7.59, 11.39]$	$IRI \in [10.80, 16.16]$	$IRI \in [16.16, 32.32]$
10 KM/H	$IRI \in [8.99, \infty)$	$IRI \in [11.39, \infty)$	$IRI \in [16.16, \infty)$	$IRI \in [32.32, \infty)$

Source: Author's calculations and Yu et al. (2006), Table 2. IRI denotes the international roughness index, measured in m/km. Ride quality levels are subjective and measured on a 5-point scale ("Very Good", "Good", "Fair", "Mediocre", and "Poor").