

# The Impact of Road Development on Household Welfare in Rural Papua New Guinea

Online Appendix

## 1. Road Summary Statistics

Table A1 shows the road lengths and conditions of PNG's road network for 2000 and 2009. Tables A2 and A3 show the transitions in surface type and condition between the two years for those segments that are included in both maps. The tables reveal no consistent trend in road development. Considering change in surface type, we observe that the length of roads upgraded (i.e., changes from gravel to sealed surface) was roughly offset by roads that deteriorated (gravel to dirt). The characterization of road condition captured in the 2000 and 2009 maps shows substantial improvement (most notably, road condition improving from poor to fair) alongside decline (mainly from good to fair). Table A4 is an extension of Table 1 in the paper, showing summary statistics of surface type *and* condition for the roads to the nearest town. It is noteworthy that for the average household, the route to the nearest town is better than the average PNG road as shown in Table 1, signified by the higher shares of sealed and gravel roads.

Table A1. Extent and surface type/condition of the main PNG road network in 2000 and 2009

Surface	Condition	2000		2009	
		Length (km)	Share	Length (km)	Share
Sealed	Good	911	7.8%	1,799	7.0%
Sealed	Fair	914	7.8%	1,067	4.2%
Sealed	Poor	314	2.7%	371	1.5%
Gravel	Good	2,137	18.3%	1,096	4.3%
Gravel	Fair	1,649	14.1%	7,300	28.6%
Gravel	Poor	4,232	36.3%	5,726	22.4%
Dirt	Good	223	1.9%	166	0.7%
Dirt	Fair	63	0.5%	3,660	14.3%
Dirt	Poor	1,230	10.5%	4,332	17.0%
All	All	11,672	100%	25,517	100%

Table A2. Transition matrix comparing road segment surface types in 2000 and 2009

	Sealed '09	Gravel '09	Dirt '09	Total
Sealed '00	1,821	226	93	2,139
Gravel '00	683	6,502	832	8,017
Dirt '00	27	304	1,185	1,516
Total	2,531	7,031	2,110	11,672

Notes: Reported in kilometres. Only listed for stretches where surface type is known in 2000.

Table A3. Transition matrix comparing road segment conditions in 2000 and 2009

	Good '09	Fair '09	Poor '09	Total
Good '00	1,077	1,531	662	3,270
Fair '00	925	970	731	2,626
Poor '00	457	2,994	2,326	5,776
Total	2,458	5,495	3,719	11,672

Notes: Reported in kilometres. Only listed for stretches where road condition is known in 2000.

Table A4. Summary statistics of road surface type and condition for the analysis sample

	1996			2009/10			p-value
	Mean	SD	N	Mean	SD	N	
<i>Impact variables (obtained from road maps)</i>							
Road to town: km of good sealed road	7.907	12.61	680	7.740	16.35	1,643	0.949
Road to town: km of fair sealed road	6.834	13.16	680	5.900	15.70	1,643	0.717
Road to town: km of poor sealed road	1.355	1.962	680	1.041	3.660	1,643	0.504
Road to town: km of good gravel road	7.315	8.773	680	2.792	8.045	1,643	0.004
Road to town: km of fair gravel road	10.53	12.78	680	11.09	18.36	1,643	0.830
Road to town: km of poor gravel road	12.62	17.20	680	3.801	10.82	1,643	0.009
Road to town: km of good dirt road	0.029	0.205	680	0	0	1,643	0.304
Road to town: km of fair dirt road	1.696	3.282	680	2.934	8.938	1,643	0.223
Road to town: km of poor dirt road	1.703	4.707	680	1.265	5.815	1,643	0.626

Notes: Road summary statistics are obtained using household sampling weights. Standard deviations and p-values account for clustering at the village level. The p-values show the probability of equal means.

## 2. Selective Migration

To check whether better roads and migration within PNG are correlated, we regress two indicators of migration on our road variables. The first variable is the ratio of prime-aged men in a household, as they are the people most likely to migrate (Aggarwal, 2018). The second variable, which is only available for the HIES 09/10, is a dummy indicating whether anyone among the household head, the spouse, the head's parents, or grandparents, has moved to the current province within the last 10 years. We include location-specific covariates and province-time fixed effects, but we do not include the Mundlak terms, as this is not a causal analysis. Table A5 shows the regression results. It appears that road quality is statistically insignificant for both migration indicators.

Table A5. Regressions of indicators of migration on road type and distances

	Ratio of men between 18 and 40	Household member moved to province ≤ 10 years ago
Distance Sealed (km)	0.0001 (0.773)	-0.0018 (0.156)
Distance Gravel (km)	0.0001 (0.864)	-0.0019 (0.104)
Total distance to nearest town (km)	-0.0001 (0.729)	0.0013 (0.124)
Distance to nearest road (km)	0.0013 (0.537)	-0.0034 (0.694)
Household controls	No	No
Mundlak terms included	No	No
R-squared	0.033	0.043
Villages	155	95
Households	2323	1643
p-value joint significance	0.958	0.231

Notes: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01. P-values in parentheses, using standard errors clustered at the village level. Both regressions are weighted using household sampling weights from both surveys. Dirt road is the excluded category for the distance variables. Road sections observed in the 2009 maps, but not in 2000, are assumed to remain the same over time. Both specifications also include location-specific control variables (see Table 1) as well as province-time-fixed effects. The last row displays the p-value of a Wald test of the joint significance of all road variables.

### **3. Impact Analysis of Surface Type and Condition**

Beside the regressions on road surface type, we also consider a more detailed set of specifications that consider surface type and road condition. Guidelines for the interpretation of the classifications of road condition are provided in the CAPE-PNG-9-Transport-Sector-Assessment. A road segment is labelled “good” if it is passable for a two-wheel vehicle in wet weather, “fair” if it requires periodic maintenance, and “poor” if it requires reconstruction or rehabilitation (<https://www.adb.org/sites/default/files/linked-documents/CAPE-PNG-9-Transport-Sector-Assessment.pdf>).

Considering both surface type and condition leaves us with nine categories of road segments (all combinations of sealed, gravel, dirt, and good, fair, poor), some of which are empty or have very few observations. To avoid including variables with little variation, we lump together road categories that accounted for less than 5% of total length. Accordingly, we combine fair and poor sealed roads into one category, and combine good, fair, and poor dirt roads into another category. The excluded category for the regressions is dirt road. The results are summarised in Tables A6 to A8, which are the equivalents of Tables 2 to 4 in the paper. Conditional on surface type, we see no significant differences in effects by road condition. This does not mean that road condition has no effect. We believe that a combination of insufficient statistical power and quality of the road condition data is responsible for this null result. Road condition can change quickly in PNG, and the way it was assessed likely varied over time and across provinces, making it a variable more prone to measurement error than surface type.

Table A6. Impact of road type on consumption and poverty status using detailed road type variables

	Log(real p.a.e. consumption)			Poverty Status		
	(1)	(2)	(3)	(1)	(2)	(3)
Sealed and good (km)	-0.0065 (0.138) [0.495]	-0.0088** (0.037) [0.366]	0.0144 (0.245) [0.603]	0.0019 (0.456) [0.786]	0.0031 (0.232) [0.533]	-0.0025 (0.731) [0.854]
Sealed and not good (km)	0.0050 (0.287) [0.659]	0.0056 (0.225) [0.533]	0.0196** (0.036) [0.541]	-0.0040 (0.112) [0.450]	-0.0043* (0.085) [0.384]	-0.0085 (0.126) [0.542]
Gravel and good (km)	-0.0031 (0.540) [0.786]	-0.0050 (0.327) [0.741]	0.0042 (0.684) [0.810]	0.0001 (0.960) [1.000]	0.0011 (0.680) [0.985]	-0.0032 (0.613) [0.734]
Gravel and fair (km)	-0.0026 (0.556) [0.786]	-0.0024 (0.559) [0.985]	0.0058 (0.609) [0.734]	0.0014 (0.536) [0.786]	0.0013 (0.558) [0.985]	-0.0037 (0.577) [0.734]
Gravel and poor (km)	-0.0028 (0.436) [0.786]	-0.0050 (0.155) [0.519]	0.0136 (0.162) [0.572]	0.0012 (0.513) [0.786]	0.0022 (0.225) [0.533]	-0.0069 (0.224) [0.597]
Total distance to nearest town (km)	-0.0021 (0.536)	-0.0018 (0.582)	0.0017 (0.689)	0.0020 (0.236)	0.0019 (0.264)	0.0005 (0.816)
Distance to nearest road (km)	0.0029 (0.873)	0.0119 (0.454)	0.0039 (0.797)	-0.0059 (0.575)	-0.0108 (0.258)	-0.0075 (0.453)
Household controls	No	Yes	Yes	No	Yes	Yes
Mundlak terms included	No	No	Yes	No	No	Yes
R-squared	0.192	0.323	0.332	0.135	0.218	0.222
Villages	155	155	155	155	155	155
Households	2153	2153	2153	2153	2153	2153
p-value Mundlak	-	-	0.003	-	-	0.009

Notes: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01. P-values in parentheses, using standard errors clustered at the village level. FDR q-values in square brackets. All regressions are weighted using person sampling weights from both surveys. Dirt road is the excluded category for the distance variables. Road sections observed in the 2009 maps, but not in 2000, are assumed to remain the same over time. All specifications also include location-specific control variables (see Table 1) as well as province-time-fixed effects. Columns labelled (2) in addition include the household-level control variables (see Table 1), and columns labelled (3) include Mundlak terms and household-level control variables. The last row displays the p-value of a Wald test of the joint significance of the Mundlak terms.

Table A7. Impact of road type on subsistence farming and wage employment using detailed road type variables

	Household does subsistence farming			Someone in the household has wage job		
	(1)	(2)	(3)	(1)	(2)	(3)
Sealed and good (km)	0.0014 (0.108) [0.450]	0.0018* (0.071) [0.384]	-0.0047 (0.183) [0.597]	0.0016 (0.360) [0.762]	-0.0003 (0.862) [1.000]	-0.0074 (0.305) [0.604]
Sealed and not good (km)	-0.0000 (0.996) [1.000]	-0.0001 (0.957) [1.000]	-0.0033 (0.354) [0.604]	0.0007 (0.707) [0.960]	0.0000 (0.979) [1.000]	-0.0118* (0.058) [0.541]
Gravel and good (km)	0.0022* (0.079) [0.403]	0.0027* (0.056) [0.383]	0.0023 (0.521) [0.734]	0.0031* (0.080) [0.403]	0.0019 (0.200) [0.533]	0.0014 (0.813) [0.910]
Gravel and fair (km)	-0.0007 (0.504) [0.786]	-0.0005 (0.614) [0.985]	0.0022 (0.540) [0.734]	0.0004 (0.800) [1.000]	0.0001 (0.972) [1.000]	0.0008 (0.870) [0.926]
Gravel and poor (km)	-0.0001 (0.952) [1.000]	0.0003 (0.754) [1.000]	0.0009 (0.778) [0.895]	0.0026 (0.159) [0.495]	0.0010 (0.523) [0.985]	0.0024 (0.612) [0.734]
Total distance to nearest town (km)	-0.0002 (0.817)	-0.0004 (0.651)	-0.0006 (0.572)	-0.0026** (0.021)	-0.0015 (0.117)	-0.0014 (0.352)
Distance to nearest road (km)	0.0078* (0.093)	0.0077* (0.085)	0.0079* (0.098)	-0.0365*** (0.000)	-0.0313*** (0.000)	-0.0358*** (0.000)
Household controls	No	Yes	Yes	No	Yes	Yes
Mundlak terms included	No	No	Yes	No	No	Yes
R-squared	0.150	0.195	0.202	0.155	0.226	0.236
Villages	155	155	155	155	155	155
Households	2323	2322	2322	2312	2311	2311
p-value Mundlak	-	-	0.004	-	-	0.800

Notes: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01. P-values in parentheses, using standard errors clustered at the village level. FDR q-values in square brackets. All regressions are weighted using household sampling weights from both surveys. Dirt road is the excluded category for the distance variables. Road sections observed in the 2009 maps, but not in 2000, are assumed to remain the same over time. All specifications also include location-specific control variables (see Table 1) as well as province-time-fixed effects. Columns labelled (2) in addition include the household-level control variables (see Table 1), and columns labelled (3) include Mundlak terms and household-level control variables. The last row displays the p-value of a Wald test of the joint significance of the Mundlak terms.

Table A8. Impact of road type on having a good roof and school enrolment using detailed road type variables

	Home has a good roof			Ratio of school children going to school		
	(1)	(2)	(3)	(1)	(2)	(3)
Sealed and good (km)	0.0033** (0.041) [0.287]	0.0026 (0.104) [0.402]	0.0064 (0.463) [0.734]	0.0031 (0.247) [0.659]	0.0006 (0.802) [1.000]	0.0163* (0.073) [0.541]
Sealed and not good (km)	0.0055** (0.022) [0.274]	0.0052** (0.028) [0.366]	0.0122 (0.132) [0.542]	0.0012 (0.627) [0.824]	0.0012 (0.611) [0.985]	0.0144* (0.087) [0.541]
Gravel and good (km)	0.0047** (0.036) [0.287]	0.0043* (0.057) [0.383]	0.0063 (0.292) [0.604]	0.0046 (0.152) [0.495]	0.0028 (0.344) [0.741]	0.0206** (0.023) [0.541]
Gravel and fair (km)	0.0043** (0.018) [0.274]	0.0042** (0.021) [0.366]	0.0087 (0.138) [0.542]	-0.0015 (0.562) [0.786]	-0.0022 (0.342) [0.741]	0.0135 (0.132) [0.542]
Gravel and poor (km)	0.0063*** (0.000) [0.010]	0.0056*** (0.001) [0.040]	0.0134** (0.015) [0.541]	0.0014 (0.599) [0.817]	-0.0004 (0.885) [1.000]	0.0145* (0.083) [0.541]
Total distance to nearest town (km)	-0.0049*** (0.001)	-0.0044*** (0.004)	-0.0036* (0.081)	-0.0013 (0.569)	-0.0001 (0.975)	0.0023 (0.186)
Distance to nearest road (km)	-0.0301*** (0.000)	-0.0267*** (0.000)	-0.0284*** (0.000)	-0.0166* (0.072)	-0.0095 (0.233)	-0.0182** (0.050)
Household controls	No	Yes	Yes	No	Yes	Yes
Mundlak terms included	No	No	Yes	No	No	Yes
R-squared	0.455	0.470	0.474	0.188	0.248	0.256
Villages	155	155	155	154	154	154
Households	2322	2321	2321	1538	1537	1537
p-value Mundlak	-	-	0.538	-	-	0.188

Notes: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01. P-values in parentheses, using standard errors clustered at the village level. FDR q-values in square brackets. The regressions of having a good roof are weighted using household sampling weights from both surveys, the ones of schooling use child sampling weights. Dirt road is the excluded category for the distance variables. Road sections observed in the 2009 maps, but not in 2000, are assumed to remain the same over time. All specifications also include location-specific control variables (see Table 1) as well as province-time-fixed effects. Columns labelled (2) in addition include the household-level control variables (see Table 1), and columns labelled (3) include Mundlak terms and household-level control variables. The last row displays the p-value of a Wald test of the joint significance of the Mundlak terms.