MATHEMATICA Policy Research

REPORT

FINAL REPORT

Evaluation of a Rural Road Rehabilitation Project in Armenia

March 12, 2015

Kenneth Fortson
Randall Blair
Kathryn Gonzalez

Submitted to:
Millennium Challenge Corporation
875 Fifteenth St., NW
Washington, DC 20005-1203
Project Officer: Rebecca Goldsmith

Submitted by:
Mathematica Policy Research
P.O. Box 2393
Princeton, NJ 08543-2393
Project Director: Kenneth Fortson
Reference Number: 06916

This page has been left blank for double-sided copying.

ACKNOWLEDGMENTS

We greatly appreciate the hard work of many people whose efforts contributed to this report. We especially thank our monitoring and evaluation colleagues: Rebecca Goldsmith, Jack Molyneaux, Sarah Bishop, Edith Petrosyan, and Emily Andrews of MCC; and, Lusine Yeremyan, Lusine Kharatyan, Shushan Kurkchiyan, and Ester Hakobyan of MCA-Armenia. Their input at all stages of the project has been invaluable. We also appreciate the cooperation and insights of the Rural Road Rehabilitation Project program manager at MCA-Armenia, Hakob Petrosyan, and our colleagues at the World Bank, especially Vickram Cuttaree and Zaruhi Tokhmakhian.

Additionally, we thank MCC's Resident Country Mission Director Alex Russin and MCA-Armenia's CEO Ara Hovsepyan. Throughout the lifespan of the Compact with Armenia, their interest in and support for M&E generally and impact evaluation particularly has been invaluable for cultivating an environment in which rigorous impact evaluation is possible and the results receive appropriate attention.

This report would not be possible without the excellent partnership between MCC, MCA-Armenia, and the National Statistical Service of Armenia. We especially thank Diana Martirosova and Armine Arushanyan. In addition to quantitative data, the present report draws on qualitative data about program implementation collected for reports by Socioscope.

Our colleagues at Mathematica have provided guidance and suggestions throughout the project, particularly Duncan Chaplin, who carefully reviewed and provided excellent comments that greatly improved the report. Jane Fortson, Phil Gleason, and Peter Schochet were instrumental in helping us develop the evaluation designs. We also thank Sheena Flowers for her diligence and patience in formatting this report. Mike Donaldson provided careful editing of the draft report. And last but definitely not least, Caitlin Matolka and Laurie Schulte have done a wonderful job tracking and managing the project staffing and resources.

All of the above organizations and individuals have contributed tremendously to this report and to the project more generally, but any remaining errors are the fault of the authors alone.

This page has been left blank for double-sided copying.

CONTENTS

This page has been left blank for double-sided copying.

TABLES

This page has been left blank for double-sided copying.

FIGURES

This page has been left blank for double-sided copying.

ABBREVIATIONS

ARD	Armenian Road Directorate
ERR	Economic rate of return
ILCS	Integrated Living Conditions Survey
MCA	Millennium Challenge Account
MCC	Millennium Challenge Corporation
LRN	Lifeline Road Network
RRRP	Rural Road Rehabilitation Project

This page has been left blank for double-sided copying.

EXECUTIVE SUMMARY

In this report, we present key findings of an impact evaluation of the Rural Road Rehabilitation Project (RRRP) in Armenia. The RRRP was originally conceived as part of a five-year, $236 million Compact between the Millennium Challenge Corporation (MCC) and Armenia designed to increase household income and reduce poverty in rural Armenia. The Compact, managed by the Millennium Challenge Account with Armenia (MCA-Armenia), included two projects: (1) the Irrigated Agriculture Project, which comprised irrigation infrastructure rehabilitation, farmer training, technical assistance to water user associations and postharvest enterprises, and access to credit for farms and agribusiness; and (2) the RRRP, which is the subject of the present report.

Background. Originally funded at $67 million, the RRRP was designed to rehabilitate up to 943 km of rural roads, or 35 percent of the government-proposed lifeline road network (LRN). These 943 km comprised 85 road links. According to the Compact, approximately 360,000 rural inhabitants in 265 rural communities would benefit from the RRRP. The project was projected to reduce transport costs for the greater rural community, including farmers and processors, by an estimated $20 million per year beginning five years after material project benefits were realized (MCC 2011). In late 2007, rehabilitation targets were reduced by over two-thirds—from 943 to 297 km—due to the dramatic devaluation of the U.S. dollar against the Armenian dram and construction price escalation (Socioscope 2010).

In 2007, construction of the pilot phase began, during which 24.4 km of roads in the marzes of Aragatsotn, Shirak, and Lori were rehabilitated with MCA funds. The original rehabilitation designs for these roads sections were completed by the Lincy Foundation. The MCA-Armenia RRRP team reviewed these designs and supervised the Armenian Road Directorate (ARD) as it paved and improved drainage systems for these pilot roads. However, in June 2009 the MCC Board made the decision to discontinue funding any further road construction and rehabilitation under the Compact due to concerns about Armenia's democratic governance. Starting in 2009, the Armenian government accessed loans from the World Bank to rehabilitate many road sections that were included in the RRRP plans before MCC discontinued funding, using and updating the road project designs developed by MCA-Armenia. From 2009 to 2011, the World Bank approved over $100 million to rehabilitate a total of 430 km of rural roads, most of which came from the original 943 km planned in the Compact. To complement these loans, the Republic of Armenia invested $16 million to rehabilitate key sections of the LRN, starting in 2009. The Armenian Road Directorate managed all World Bank- and Armenian government-funded rehabilitation efforts, with assistance from international consulting firms and World Bank technical staff. As of December 2013, 446 km of roads had been improved with World Bank funds (World Bank 2013), and approximately 50 km of roads had been improved with Armenian government funds.

The Armenian government initially received a $25 million World Bank loan to rehabilitate road sections. In August 2009, the World Bank approved additional financing of $36.6 million in loans to rehabilitate another 146 km of LRN, and in 2010, the World Bank approved another $40 million in low-interest loans. This brought the Bank's total lending for road rehabilitation of over $100 million to rehabilitate a total of 430 km of rural roads from the original 943 km planned in the Compact. As illustrated in Table I.1, the final MCA-funded portion of the RRRP was $8.4

million, compared to the final World Bank–funded allocation of over $100 million and the government of Armenia's contribution of $16 million.

Table 1 summarizes the planned investments by MCA-Armenia and the investments ultimately made by MCA-Armenia and the World Bank, and Figure 1 summarizes the RRRP timeline. As illustrated, completed World Bank–funded rehabilitation of 446 km of rural roads far exceeds the final MCA-funded rehabilitation of 24.4 km of roads, but is less than half of the original target of 943 km of rehabilitated roads cited in the Compact. Although most road rehabilitation was ultimately funded by the World Bank, from this point forward we refer to all road links that were originally planned to be rehabilitated by MCA-Armenia as part of the RRRP for expositional simplicity.

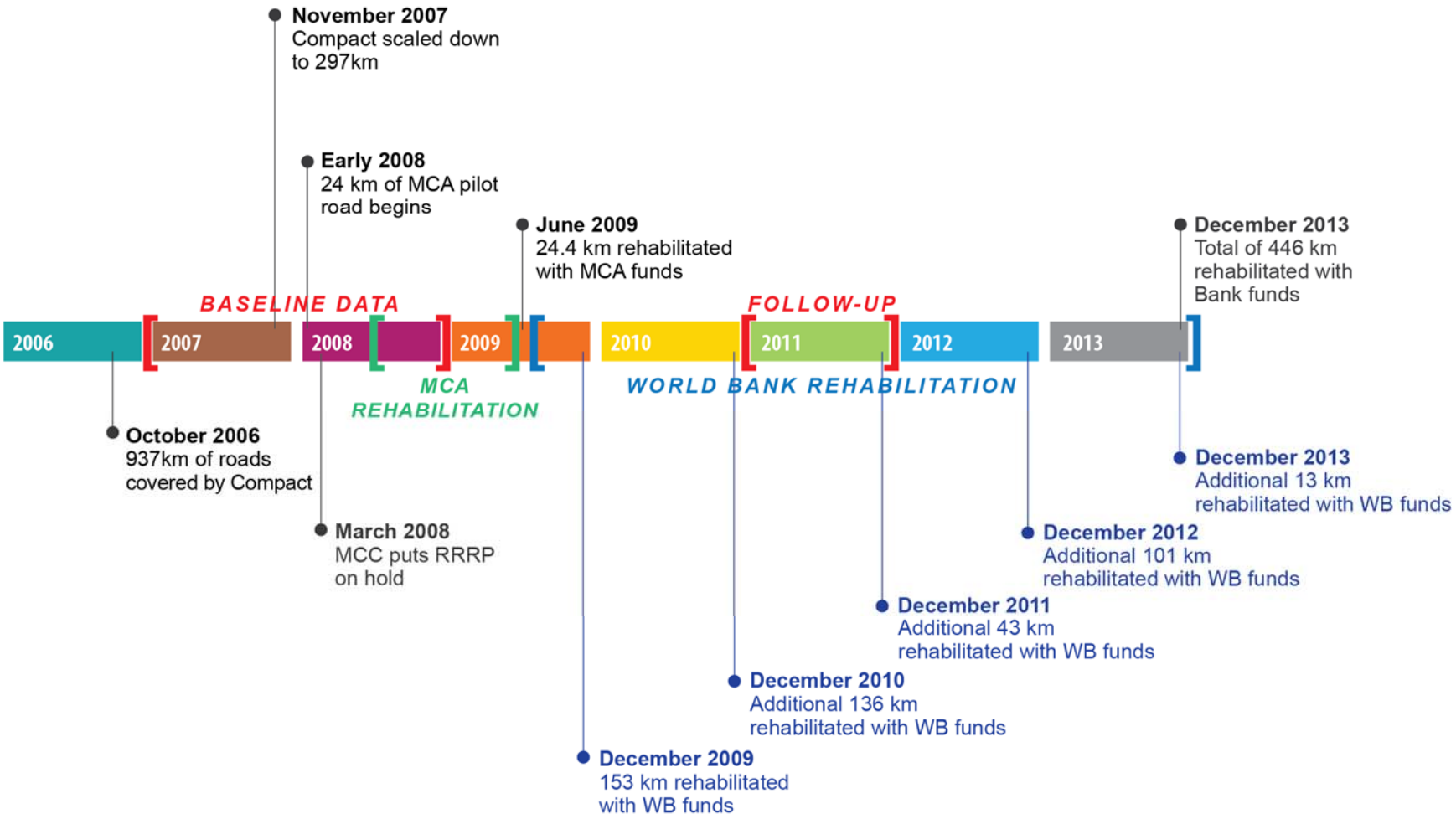
Table 1. Summary of RRRP investments and targets in Armenia

	Compact	Rescoped MCA Project (2008)	Final MCA Project (2009)	Armenian Government/World Bank Project (2008–2013)
Funding (in U.S. millions)	$67	$67	$8.4 in MCA funds	$101.6 in World Bank funds; $16 in RA funds
Rehabilitation target (km)	Up to 943	297	--	430
Rehabilitation completed (km)	--	--	24.4	446

Sources: MCC-Armenia Compact 2007, Socioscope 2010: Rural Roads Rehabilitation Project Qualitative Process Analysis Final Report, and World Bank 2014.

RA = Republic of Armenia

Figure 1. Timeline of road construction and data collection

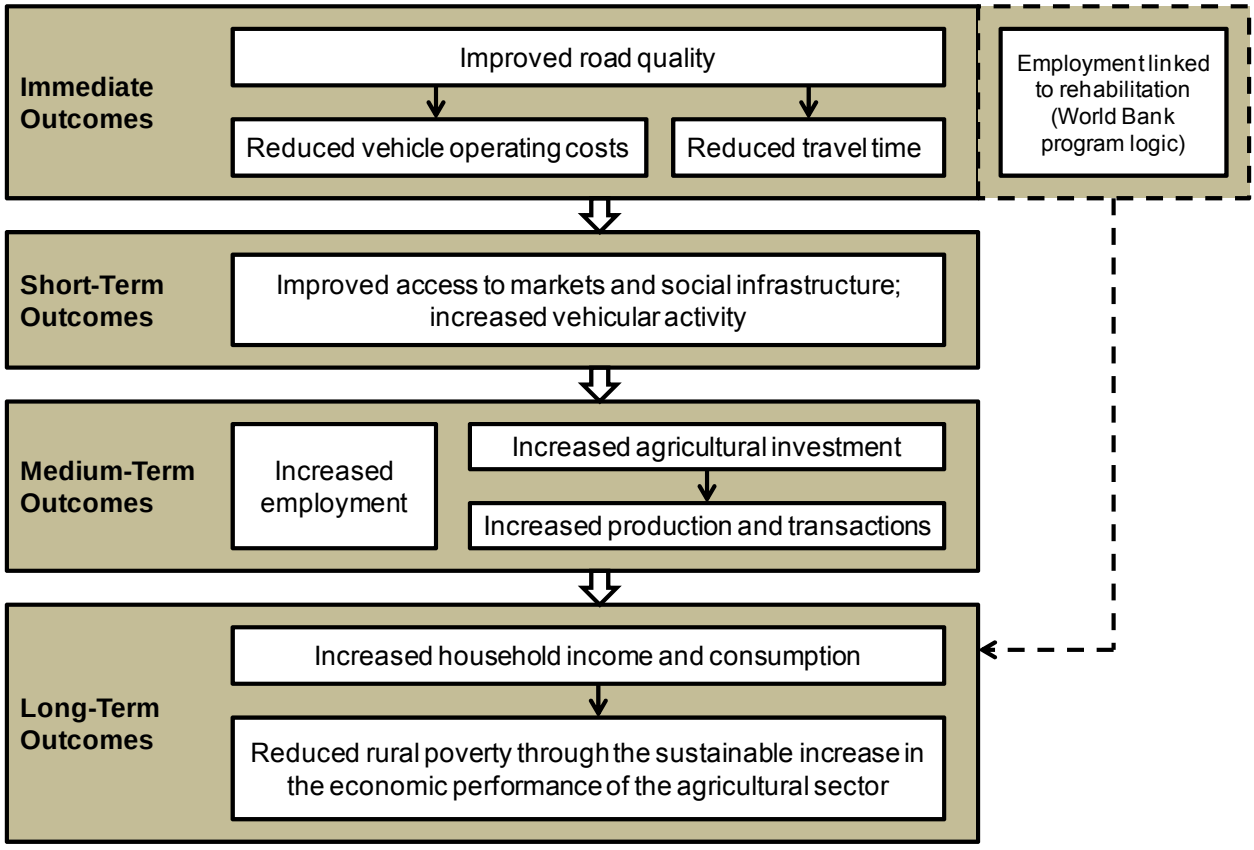


RRRP Program Logic. Figure 2 illustrates how the RRRP was originally conceived to contribute to the Compact's ultimate goal of reduced poverty. Improved rural roads would reduce travel time and vehicle operation costs, which would enhance residents' access to markets and social infrastructure. Due to lower transport costs and reduced travel time, residents could access inputs at cheaper prices and potentially increase their agricultural production. Also due to improved roads, an increased number of retailers and buyers of agricultural products could access the communities, thus creating conditions for farmers to sell a larger share of their agricultural production—potentially at a better price if improved access fostered competition among buyers. Better road infrastructure could also result in non-agricultural employment opportunities for residents, which would improve household income and consumption, and decrease poverty rates in the long-term.

The RRRP took place in the context of the world financial crisis, which had a major impact on Armenia. In this context, the World Bank–financed road rehabilitation program had a stronger focus on short-term job creation than the Compact-funded project.[1] The program logic for the World Bank–financed rehabilitation of rural roads was mostly similar to Compact program logic, but because of poor economic conditions at the time, one important difference was the World Bank's explicit objective of generating temporary employment related to road construction contracts as an immediate outcome of the rehabilitation work. This additional component of the program logic is represented by the right-hand box in the Immediate Outcomes section of Figure 2.

[1] Short-term jobs associated with rehabilitation efforts were not a target outcome of the MCC-funded RRRP. The potential benefits of short-term job creation do not factor into MCC's economic return models, as these short-term jobs are considered a transfer of resources from public institutions to workers rather than economic growth. Also, MCC would not consider those that obtain short-term infrastructure employment as project beneficiaries.

Figure 2. Logic model for roads investments



A. Evaluation questions and methodology

The original scope of the impact evaluation was to empirically test whether intermediate and long-term outcomes outlined in the program logic were realized for the Compact-funded RRRP. This evaluation scope was no longer relevant once MCC discontinued funding, but the question of whether road rehabilitation had impacts could be examined using the data that were being collected for the impact evaluation. Because the fundamental research question was unchanged, and because MCC had already committed to fund the data that would be used, MCC decided to proceed with funding the evaluation of the RRRP. Due to changes in project ownership and funding, however, the central objective of this evaluation is now to determine the impact of MCA-financed road rehabilitation designs that were implemented with World Bank funds.

It should be emphasized that the road projects were completed later than they would have been had Compact funding continued for the RRRP. Consequently, the full medium- and longer-term impacts envisioned under the original program logic are unlikely to have materialized by the follow-up. Although we would not expect impacts on longer-term outcomes to have fully materialized in the time frame of the evaluation, these are central to the ultimate goals of the RRRP, and household income and its components are therefore still included in our analysis to assess if there is evidence of early impacts.

Drawing on the program logic in Figure 2, the key research questions guiding our impact evaluation of the RRRP are as follows:

- ***Did rehabilitating roads affect the quality of roads?*** We would expect road quality to be affected immediately and reflected in residents' assessments of the quality of regional roads.
- ***Did rehabilitating roads improve access to markets and social infrastructure?*** Several indicators allow us to measure access, including residents' accounts of market access; the time it takes to travel to hospitals and schools; and the availability of transportation. Access is hypothesized to be impacted in the short term.
- ***Did rehabilitating roads improve income from employment?*** Short-term construction employment was a key objective of the World Bank and is hypothesized to be immediately observable. Improved roads are also hypothesized to increase employment income in the longer term by increasing access to job opportunities, but this effect is unlikely have fully materialized in our follow-up time frame.
- ***Did rehabilitating roads affect agricultural productivity and profits, and if so, by how much?*** We measure potential effects of the RRRP on the volume and variety of agricultural and food production, as well as income from agricultural and food sales. Impacts on these outcomes are not expected to fully materialize until a few years after roads are rehabilitated, and with the short follow-up period of the present report, we would not expect large impacts but might observe early improvements.
- ***Did rehabilitating roads improve household well-being for communities served by these roads, especially income and poverty?*** Likewise, these outcomes are only expected to fully manifest in the medium and long term, but we could potentially observe early impacts.

Integrated Living Conditions Survey. The data for the RRRP impact evaluation come from the Integrated Living Conditions Survey (ILCS). The ILCS is an annual, nationally representative household survey fielded by the National Statistical Service of Armenia. Through the Compact, MCC funded an increase in the sample size and a longer survey questionnaire from 2007 to 2011 to facilitate the RRRP evaluation. During these years, the ILCS oversampled communities served by roads in MCA's initial set of rehabilitation-eligible roads—that is, the communities that were directly connected to the LRN by the specified road. The total sample included approximately 2,200 households served by 82 of the 85 original eligible roads during each year of the ILCS, though as discussed below, only a subset are included in our analysis. The standard ILCS questionnaire was also expanded during the evaluation period to include questions about intermediate outcomes related to the RRRP.

Evaluation Design and Sample. A crucial objective of any impact evaluation is to assess not only how key outcomes for the treatment group—the affected beneficiaries of the RRRP in this case—changed after the intervention, but also to assess the counterfactual: how outcomes would have changed even in the absence of the intervention. We used a difference-in-differences framework to analyze RRRP impacts. The central idea behind difference-in-differences estimation is to estimate how outcomes changed for communities served by treatment roads before and after road rehabilitation (the first difference) and compare this to how outcomes changed for communities served by comparison roads over the same time period (the second

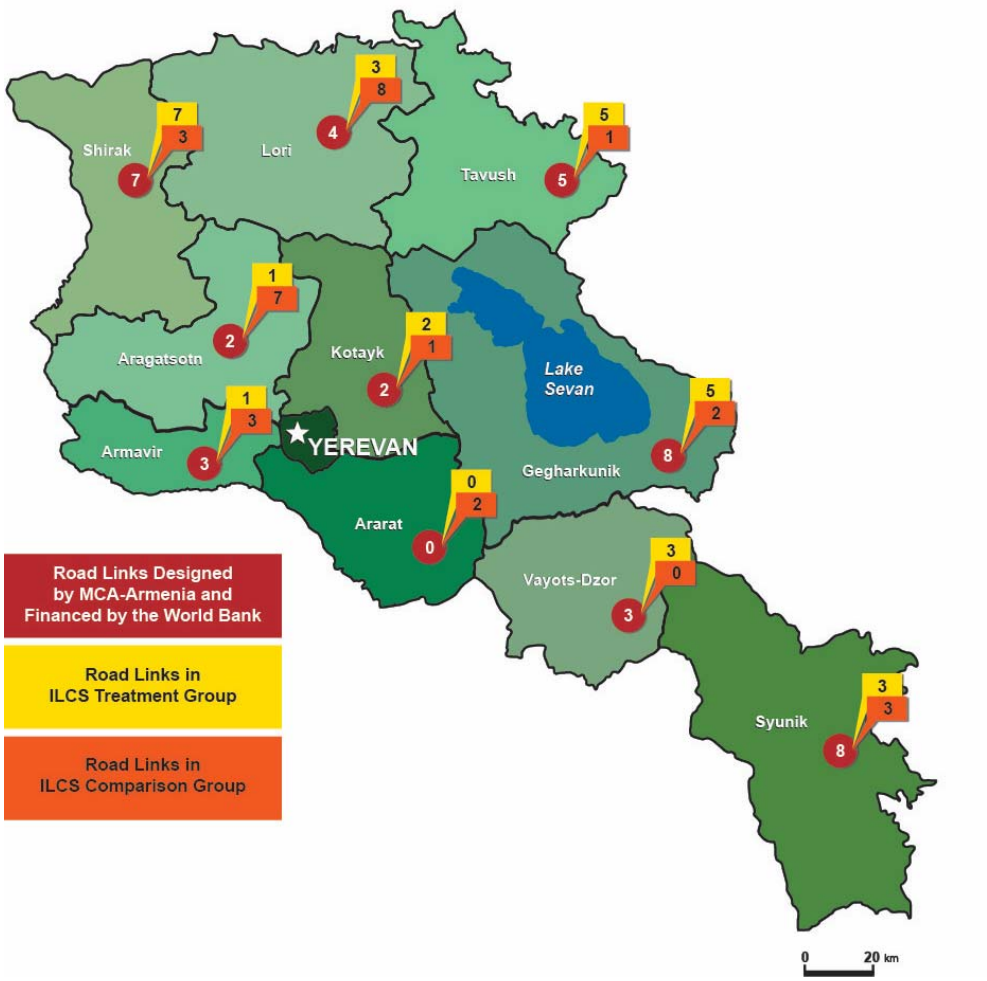
difference). The difference between these two differences is the estimated RRRP impact for these roads.

The treatment group in our analysis comprises communities served by the 27 road links that were in MCA-Armenia's original RRRP plans, were ultimately rehabilitated in 2009 and 2010, and were covered by the ILCS from 2007/2008 to 2011. Unfortunately, ILCS data were unavailable for the relatively small number of communities affected by MCA-funded rehabilitation of the pilot roads. As a result, all 27 treatment links in the evaluation sample were funded with World Bank loans to Armenia. The comparison group in our analysis comprises communities served by the 28 road links that were in MCA-Armenia's original RRRP plans, were not ultimately rehabilitated with financing from the World Bank as of 2011, and were covered by the ILCS from 2007/2008 to 2011. Our data include outcomes measured before any roads were rehabilitated and after they were completed for both the treatment and the comparison group road links. All told, our analysis sample includes over 50 communities and 2,300 households in each of the treatment and comparison groups, for a grand total of 107 communities (55 in treatment and 52 in comparison) and 4,848 households (2,560 in treatment and 2,288 in comparison). However, the ILCS is a repeated cross-sectional survey, with different households sampled within communities from one round to the next.

As illustrated in Figure 3, road projects and the ILCS sample coverage span most of Armenia. All marzes,[2] with the exception of Ararat, had at least one road link designed by MCA-Armenia and financed by the World Bank. Similarly, all marzes have at least one treatment or comparison road link in the evaluation, and all but two marzes have at least one road link in the treatment group and one in the comparison group.

[2] Armenia is divided into 11 administrative divisions. Of these, 10 are marzes, or provinces, and the country's capital, Yerevan, is granted special administrative status.

Figure 3. Distribution of treatment and comparison road links



Design Limitations. A primary limitation of the difference-in-differences estimation approach described above is that there may be systematic differences between treatment and comparison communities—besides access to rehabilitated roads—that could affect key outcomes in the evaluation. To the extent that the comparison group differs from the treatment group along dimensions important for outcomes, we are unable to distinguish actual program impacts from underlying differences in the two groups unless we can credibly identify and control for these preexisting differences. As van de Walle (2009) notes, several factors that we do not observe in the data but could bias impact estimates in comparison studies such as ours include political power, local leadership, social capital, and social empowerment. Additionally, the estimation strategy accounts for preexisting differences in baseline outcomes, but not preexisting differences in the trends of those outcomes. We unfortunately do not have sufficient pre-rehabilitation data to test whether the treatment and comparison communities had similar growth rates for the key outcomes, such as household income and poverty.

Another limitation of the impact analysis is that our estimates will not be very statistically precise because of features of the evaluation design. Relative to an unclustered sample design, the clustered nature of the intervention reduces statistical power. Limited statistical power means that we will not be able to detect small impacts, as we might expect at the one-year follow-up (when full impacts have not yet been realized).

B. Findings from implementation analyses

Compact-funded rehabilitation was considered a success. In addition to the impact estimates based on the ILCS, MCA-Armenia commissioned an independent qualitative process analysis that examined the RRRP's design and implementation based on in-depth interviews with a smaller set of stakeholders, including program managers, construction firms, and community residents. Overall, the qualitative process analysis concluded that most of the outcome targets set for the RRRP in the MCA-Armenia Monitoring and Evaluation Plan had been fully met on the 24.4 km of pilot road sections that were rehabilitated. Particularly, as a result of project implementation, road roughness decreased, transportation costs decreased, and vehicular activity increased. The report also concluded that most of the activities envisioned under the RRRP had been effectively implemented; there were no major delays in project implementation, and the construction outputs were largely achieved (Socioscope 2010).

World Bank-funded rehabilitation was also considered successful. Regarding the implementation of the World Bank–funded road rehabilitation project, program administrators judged the project to be a success (World Bank 2010). A document that summarized all World Bank-funded rehabilitation in 2009 stated that several key factors led to the successful implementation of the project. These included the Armenian government's full support for the project, a competent implementation team from the Armenian Road Directorate, preexisting investments in selecting road links and completing construction designs, technical supervision provided by international firms, and local capacity and willingness to try new designs and technologies (World Bank 2010).

The World Bank concluded that the project dramatically surpassed the target of a 20 percent reduction in travel time on rehabilitated road links, with an average decrease of 59 percentage points in travel time across all links. In addition, the RRRP resulted in an average reduction in transport costs of 26 percent, based on measured roughness values. The World Bank also tracked the contribution of the roads rehabilitation projects to temporary employment, as this was a key objective of its investment. The World Bank estimated that the project generated nearly 40,000 person-months of employment from 2008 to 2013 (World Bank 2014).

C. Estimated RRRP impacts on perceived quality, market access, utilization, and access to social infrastructure

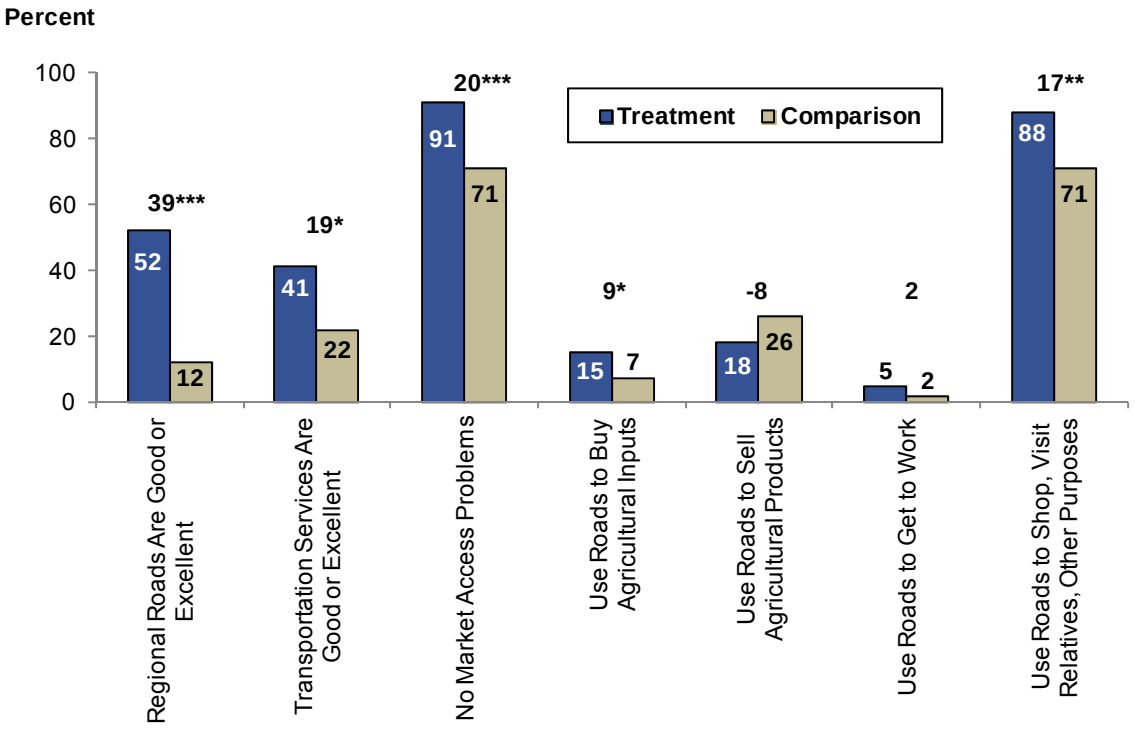
Beneficiary households perceived the improved road quality and associated transportation quality. We found statistically significant impacts on households' perceptions of road quality and their use of road links (Figure 4). Treatment households were 39 percentage points more likely than households in the comparison group to rate regional roads as good or excellent. We also observed positive, statistically significant impacts on the likelihood of households reporting transportation services as good or excellent.

The impact of road rehabilitation on reported problems with market access was also large. Treatment households were 20 percentage points more likely than households in the comparison group to report no problems with market access. Treatment group households were also more likely to report that transportation services were good (41 percent versus 22 percent in the comparison group). We also observed a positive and statistically significant impact on the likelihood of utilizing roads for a purpose other than getting to work or buying and selling

agricultural products (17 percentage points). These "other" purposes typically included shopping or visiting relatives.

Beneficiary households were more likely to use roads to buy agricultural inputs but spent less time on roads to sell their production. Although treatment group households were no more likely to use roads to sell agricultural products or for getting to work, they were nine percentage points more likely than comparison group households to use roads to buy agricultural inputs (Figure 4). In addition, households near rehabilitated roads spent less time on roads devoted to selling their agricultural production (treatment-comparison difference of -0.6 days per month, not shown). This last finding may signal the increased use of roads by agricultural intermediaries, which may have traveled directly to households to buy their production in bulk as a result of recent road rehabilitation. Another possibility is that rehabilitated roads cut down on travel time, to the extent that treatment households spent significantly less time in transit.

Figure 4. Impacts on perceived road quality and utilization (percentage of respondents reporting)



Source: 2007, 2008, and 2011 ILCS.

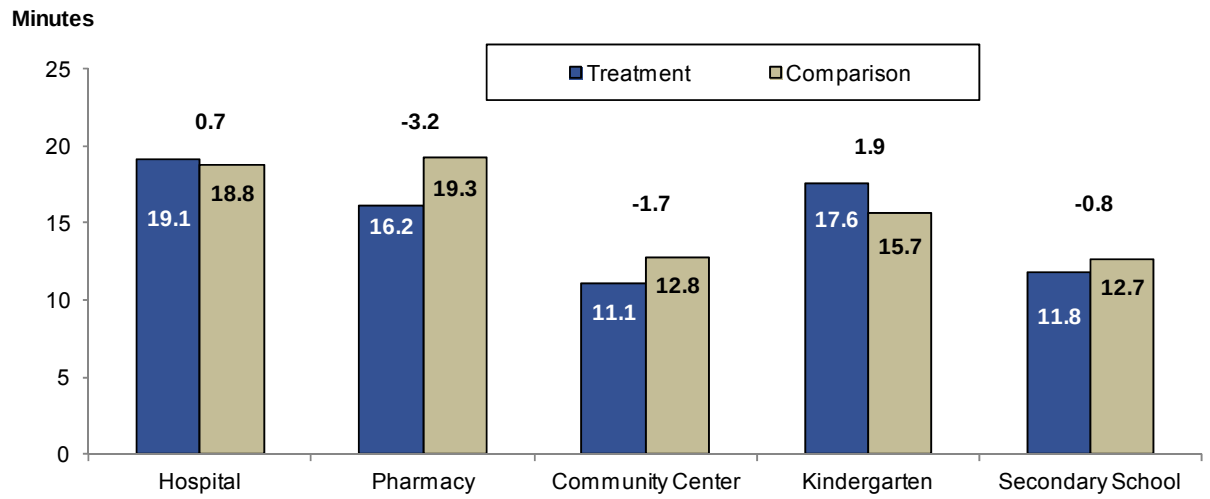
Note: All dollars are in 2011 dollars. Impact estimate may not equal difference in treatment and comparison group means due to rounding.

*/**/*** statistically significant at 10, 5, and 1 percent, respectively

Households did not report that social infrastructure was more accessible. We found no statistically significant impacts on short-term outcomes of access to social infrastructure and utilization. Residents' average reported travel time to the hospital, pharmacy, community center, and schools was similar among treatment and comparison group households (Figure 5). We note

that many community centers, kindergartens, and secondary schools are located within the communities and that it is less likely that households travel to them by the type of regional roads that were rehabilitated. Households often must travel outside of their communities to hospitals and pharmacies, but because these are less frequent trips, households may not yet be aware of any changes in travel time to hospitals or pharmacies.

Figure 5. Impacts on access to social infrastructure (average distance in minutes)



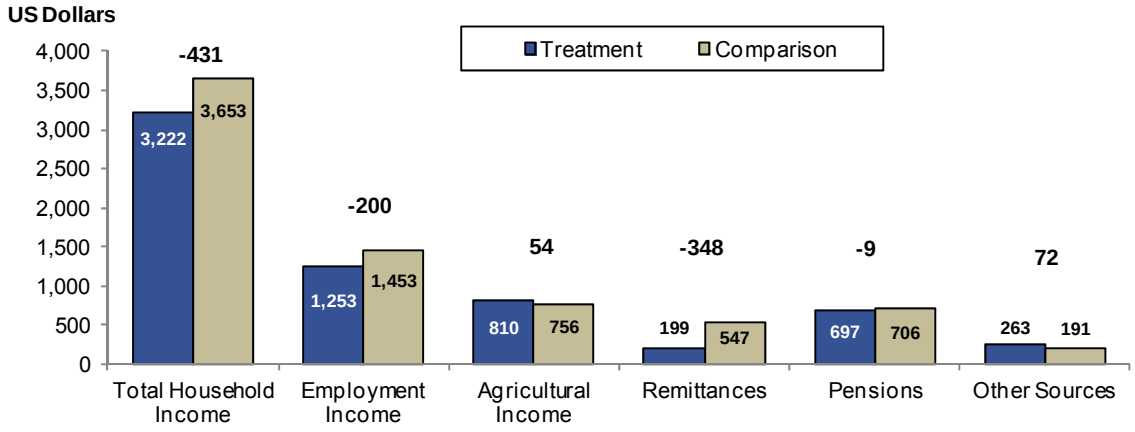
Source: 2007, 2008, and 2011 ILCS.

Note: All dollars are in 2011 dollars. Impact estimate may not equal difference in treatment and comparison group means due to rounding.

D. Estimated impacts on household income and poverty

There were no observed impacts on income within the short evaluation time frame. Total household 2011 income was not significantly different between treatment and comparison households, nor were there impacts on specific sources of income (Figure 6). Though the impact estimates are all statistically insignificant, the magnitudes of some estimates are large. In particular, the estimated impact of −$424 on total household income is 12 percent of the comparison households' average income of $3,622, and the estimated impact of −$328 on remittances is about 60 percent of the comparison households' average remittances. It should be noted that income is highly variable in the data, especially income from remittances, and only large impacts on income can be reliably detected. Because of the short follow-up period for this evaluation, impacts on income are not a priority of the present report, but the reality of imprecise impact estimates for household income should be borne in mind if a future evaluation examines longer-term impacts.

Figure 6. Impacts on 2011 income (2011 USD)



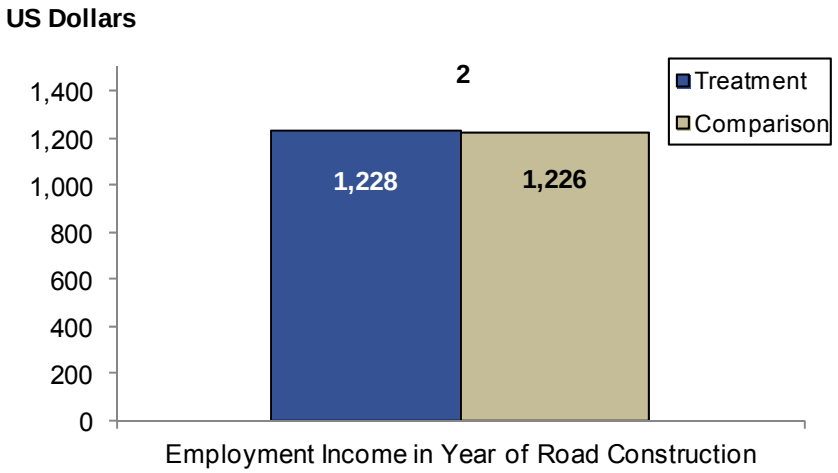
Source: 2007, 2008, and 2011 ILCS.

Note: All dollars are in 2011 dollars. Impact estimate may not equal difference in treatment and comparison group means due to rounding.

There were also no impacts on short-term employment income. As described above in the program logic, one key goal of the World Bank–funded portion of the RRRP was to increase short-term employment by employing local workers in the rehabilitation projects. Thus, we also examined impacts of the RRRP on employment income in the years in which projects were implemented. Using our main regression specification and the full comparison group of 52 villages that did not receive rehabilitation, we find no evidence of impacts on short-term employment (Figure 7). We note, however, that whereas nearly of the other impact estimates were robust to alternative empirical specifications, the estimated impacts on short-term employment were not. In one of the five alternative specifications—in which we reduce the treatment and comparison groups to only those villages with a similar range of initial economic rates of return (ERRs) for road rehabilitation projects—the estimated impact was positive and marginally statistically significant. Our evidence on possible impacts on short-term employment is less clear than for other outcomes.

There are several possible explanations for a possibly null finding on short-term employment. One is that household members employed in rehabilitation projects substituted this construction work for other employment that they would have otherwise had, including working abroad in Russia, with no change in their net income. Another possible explanation is that employment generated by the RRRP, estimated as 10,000 person-months during this time frame, was small relative to the population in these communities, thus diluting the per-household impact that we measure using ILCS data. A third possible explanation is that many of the households whose members were employed by the RRRP do not reside in the communities that were served directly by the RRRP, and thus are not covered by the ILCS analysis sample. All of these explanations may contribute to some degree, but we unfortunately cannot distinguish between them or assess their relative contributions to the null impact estimate.

Figure 7. Impact on short-term employment income (2011 USD)

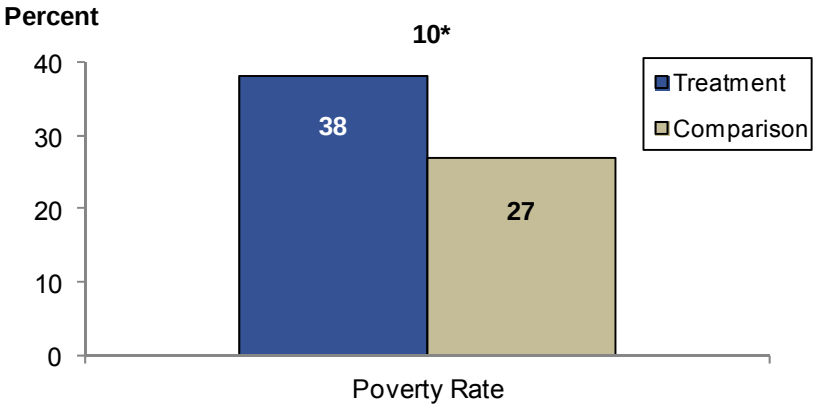


Source: 2007, 2008, 2009, 2010, and 2011 ILCS.

Note: All dollars are in 2011 dollars. Impact estimate may not equal difference in treatment and comparison group means due to rounding.

We estimated a significant and positive impact on poverty. Treatment households were 9 percentage points more likely to be poor than the comparison households; this estimate is statistically significant at the 10 percent level (Figure 8), though in our sensitivity analyses it was not always statistically significant. This non-negative (that is, detrimental) impact on poverty is concentrated among male-headed households, on which the impact on poverty is 12 percentage points and is significant at the 5 percent level (not shown). The detrimental estimated impact on poverty is a result for which we do not have a compelling explanation; we suspect that it is an anomaly due to chance differences in the sample.

Figure 8. Impact on poverty rates



Source: 2007, 2008, and 2011 ILCS.

Note: All dollars are in 2011 dollars. Impact estimate may not equal difference in treatment and comparison group means due to rounding.

*/**/*** statistically significant at 10, 5, and 1 percent, respectively

E. Summary of Findings and Policy Implications

To summarize the findings, Table 3 revisits the program logic that was illustrated in Figure 2. We found large impacts in several immediate and short-term outcomes of improved road quality and market access, but no evidence of impacts on medium- or longer-term economic outcomes. This was expected to some extent, given the time frame of the evaluation—between one and two years following road rehabilitation—as well as the analysis's lack of statistical precision. However, there were not impacts on some behaviors that might indicate large impacts on income could be expected. For example, the analysis revealed that treatment households were not any more likely to increase their agricultural investments or change their employment patterns, outcomes that we could have expected to observe within the time frame of the evaluation.

Table 2. Evidence assessment for impact of RRRP investments

Outcomes in Program Logic		Evidence Assessment
Immediate	Improved road quality	Strong evidence of large impacts (39 percentage point increase in favorability rating of regional roads)
	Reduced vehicle operating costs	Strong indirect evidence of large impacts (19 percentage point increase in approval for transportation services; 17 percentage point increase in use of roads for noncommercial purposes; decrease in time spent using roads to sell agricultural production)
	Reduced travel time	
	Nonpermanent employment linked to construction	Inconclusive; main specification finds no evidence, but alternative specifications reveal possible positive impacts
	Improved access to markets	Strong evidence of large impacts in market access (20 percentage point decrease in market access difficulties)
Short term	Improved access to social infrastructure	No evidence of impacts
	Increased vehicular activity: commercial	Some evidence of impact (increase in use of roads to buy agricultural inputs and decrease in days roads were used to sell agricultural production)
	Increased vehicular activity: non-commercial	Strong evidence of large impacts (17 percentage point increase in use of roads for noncommercial purposes);
Medium term	Increased investment	Limited evidence of small impacts (1.5 more sheep owned by treatment households and $31 increase in annual animal purchases)
	Increased employment	No evidence of impacts
	Increased production	Limited evidence that may be anomalous (5 percentage point increase in jam production, 19 kg increase in preserved vegetable production, but $31 decrease in egg sales)
	Increased transactions	No evidence of impacts
Long term	Increased household income	No evidence of impacts
	Increased household consumption	No evidence of impacts
	Reduced rural poverty	Some evidence of increase in rural poverty; likely an anomaly due to sample composition

It is difficult to assess the policy implications of these findings after the relatively short follow-up period of this evaluation, given that it is unclear if road rehabilitation efforts will eventually generate the desired outcomes of increased agricultural productivity and household income in future years, especially when coupled with the imprecision of some of the impact

estimates for key outcomes such as household income. Thus, we can do little but speculate on the policy implications. However, one sound implication of these findings is that road rehabilitation efforts appear capable of altering households' activities and use of transportation services in the short term, but they probably were not sufficient to stimulate agricultural production and sales in a time frame of one to two years.

This page has been left blank for double-sided copying.

I. BACKGROUND AND METHODS

A. Overview of the Compact and the Rural Road Rehabilitation Project

The Republic of Armenia was left with the legacy of a centrally planned economy when it declared independence from the Soviet Union in 1991. The Armenian economy was highly dependent on its Soviet trading partners and poorly equipped to function with the lack of infrastructure investment and support after Soviet withdrawal. In 1994, the Armenian government adopted a comprehensive stabilization and reform program that dramatically lowered inflation and led to steady economic growth beginning in 1995. Evidence from the Integrated Living Conditions Survey (ILCS), however, suggests that this growth occurred primarily in urban areas. As of 2004, the poverty rate in rural areas was 32 percent (National Statistical Service 2010).

The aim of the Millennium Challenge Corporation's Compact with Armenia ("the Compact"), a five-year agreement signed in March 2006, was to increase household income and reduce poverty in rural Armenia through improved performance of the country's agricultural sector. The Compact, managed by the Millennium Challenge Account with Armenia (MCA-Armenia), was originally designed to include two projects: (1) the Irrigated Agriculture Project and (2) the Rehabilitation of Rural Roads Project (RRRP). The RRRP is the subject of the present report, but because it was designed to work in tandem with the Irrigated Agriculture Project as part of the Compact, we briefly discuss both sets of activities for context.

The Irrigated Agriculture Project comprised two complementary activities—the Infrastructure Activity, through which irrigation infrastructure would be rehabilitated, and the Water-to-Market (WtM) Activity, which would provide training, technical assistance, and access to credit for farms and agribusiness. WtM was intended to help farmers harness the improvements in irrigation to introduce new technologies and shift to production of high-value agricultural crops, both of which would increase their annual income.

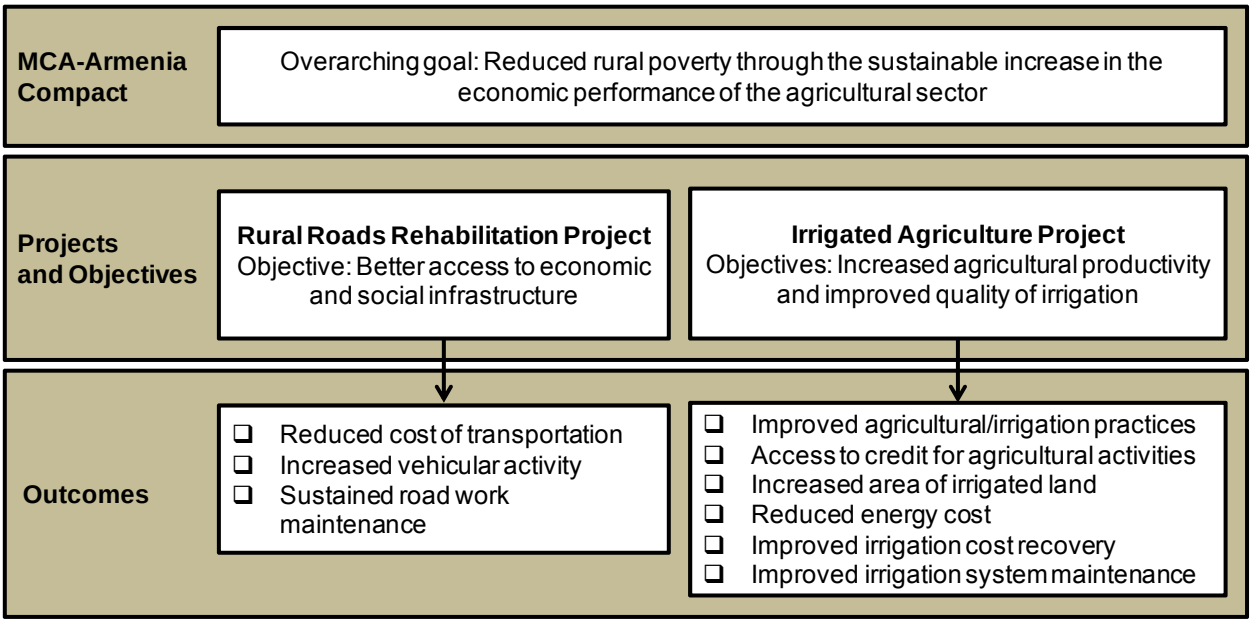
Originally funded at $67 million,[3] the RRRP was designed to rehabilitate up to 943 km of rural roads, or 35 percent of the government-proposed lifeline road network (LRN). These 943 km comprised 85 road links. According to the Compact, approximately 360,000 rural inhabitants in 265 rural communities that were directly connected to the LRN by these 85 road links would benefit from the RRRP. The project was projected to reduce transport costs for the greater rural community, including farmers and processors, by an estimated $20 million per year beginning five years after material project benefits were realized (MCC 2011).

The Irrigated Agriculture Project and the RRRP were designed as complementary investments. Farmers that participated in WtM training and benefited from improved irrigation infrastructure could use improved rural roads (funded by the RRRP) to transport their increased and more diversified agricultural production to markets and sell it for a profit. In this sense, WtM training, irrigation investments, and road investments were all oriented toward the same final objective of decreased rural poverty through improved agricultural production and sales. Figure

[3] All monetary values are in U.S. dollars except when noted otherwise.

I.1 summarizes the overall goal of the Compact and the key outcomes for each large-scale project.

Figure I.1. Overview of the Compact with Armenia



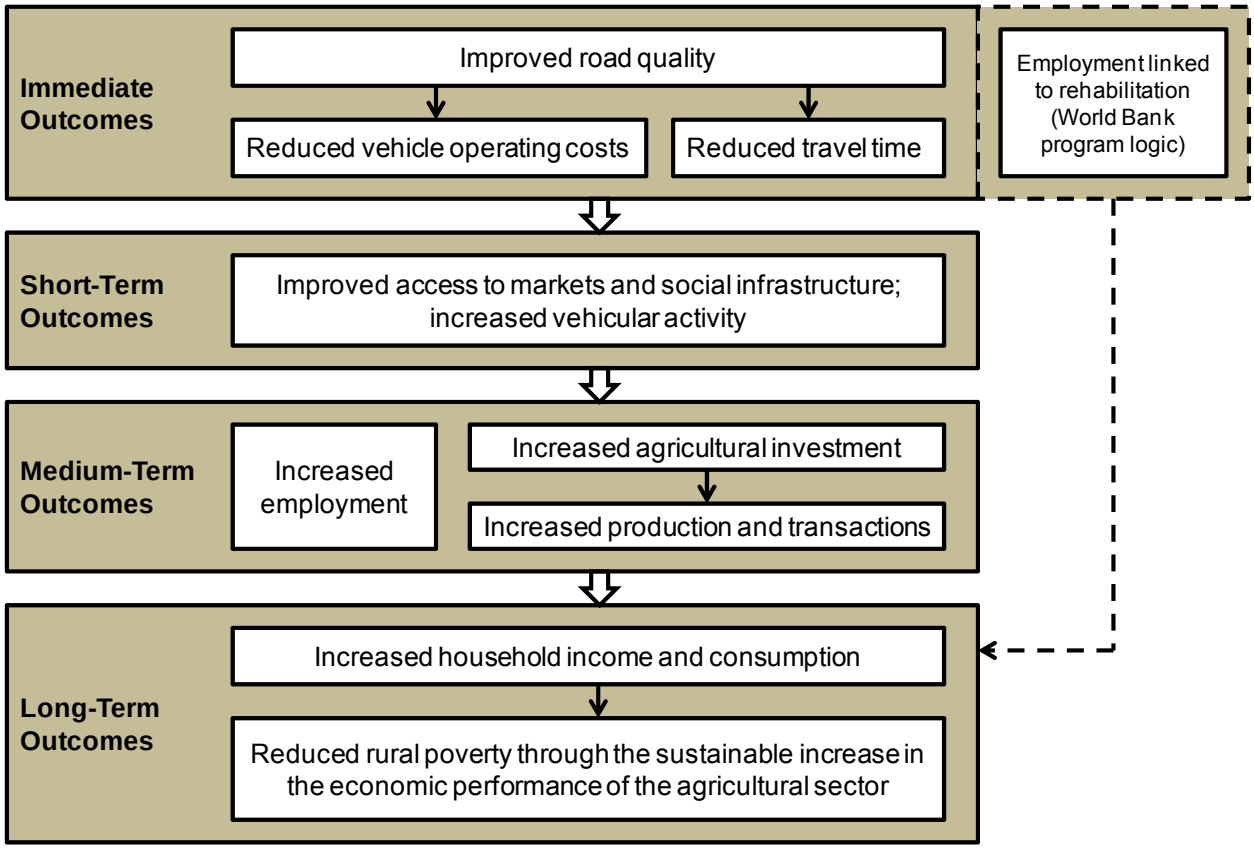
The Millennium Challenge Corporation (MCC) has commissioned evaluations to examine the RRRP, the Infrastructure Activity, and the WtM Activity. The goal of these evaluations is to determine the extent to which these projects and activities realized their objectives, and to document lessons that could be applied to future infrastructure and agriculture interventions. This report focuses on the evaluation of the RRRP. Mathematica has also produced a separate report on the evaluation of the WtM Activity (Fortson et al. 2013) and will submit an additional impact report on the Irrigation Infrastructure Activity. However, whenever relevant, each of these reports discusses planned and actual interactions between and within these large-scale projects.

Figure I.2 illustrates how the RRRP was designed to contribute to the Compact's ultimate goal of reduced poverty. Improved rural roads would reduce travel time and vehicle operation costs, which would enhance residents' access to markets and social infrastructure. Due to lower transport costs and reduced travel time, residents could access inputs at cheaper prices and potentially increase their agricultural production. Also due to improved roads, an increased number of retailers and buyers of agricultural products could access the communities, thus creating conditions for farmers to sell a larger share of their agricultural production—potentially at a better price if improved access fostered competition among buyers. Better road infrastructure could also result in non-agricultural employment opportunities for residents, which would improve household income and consumption, and decrease poverty rates in the long-term.

As discussed later, this program logic was modified under the World Bank–funded portion of the project, which identified increased employment linked to rehabilitation efforts as a key

immediate outcome of the project (represented in the right-hand box of the Immediate Outcomes section in Figure I.2).

Figure I.2. Logic model for roads investments



The impetus for the RRRP was the government-proposed lifeline road network (LRN), which was conceived as a network of roads spanning nearly 3,000 km that would provide every rural community with essential road access to markets, social services, and the main road/interstate network (World Bank 2010)[4] From these, the RRRP selected LRN road links that were identified as highly cost-effective according to estimated economic rates of return (ERRs). During Compact development proceedings, stakeholders identified 943 of the LRN's 3,000 kilometers for inclusion in the RRRP.[5]

[4] The LRN comprises large national roads and about 2,250 km of local roads that were reclassified as national (or republican) roads in 2008. The lifeline road concept identifies a priority network ensuring at least one access road for all 960 communities in Armenia. Approximately 58 percent of these roads carry traffic higher than 300 vehicles per day and 20 percent carry traffic higher than 1,000 vehicles per day.

[5] This includes 321 km of national roads and 622 km of local roads.

In late 2007, rehabilitation targets were reduced by two-thirds due to the dramatic devaluation of the U.S. dollar against the Armenian dram and construction price escalation. Using new ERR calculations based on recent feasibility studies, stakeholders selected 297 km of rural roads for rehabilitation under the RRRP. Rehabilitation efforts were to be implemented with one pilot phase and three separate packages. In all, planned construction was designed to benefit 84 communities and 142,909 rural residents (Socioscope 2010).

In 2007, construction of the pilot phase of the RRRP began. During this phase, two road sections spanning a total of 24.4 kilometers were slated for rehabilitation. A portion of these 24.4 kilometers overlapped with the 943 kilometers of the LRN that had been slated for the RRRP. The project was managed by the MCA-Armenia RRRP team and implemented by the Armenian Road Directorate (ARD), the government ministry that handled all construction and maintenance of the national road system. Several companies also provided feasibility/design studies and construction supervision, including the Swedish firm Sweco International and a consortium of the German firm Kocks Consult GmbH and the Armenian firm Arosa Ltd. Major project activities included road leveling and repaving, as well as technical assistance to construction firms. Rehabilitation of these two road sections was successfully completed in June 2009.

However, at the June 2009 MCC Board meeting, the decision was made not to continue funding any further road construction and rehabilitation under the $236 million Compact due to concerns about Armenia's democratic governance. The RRRP was originally put on indefinite hold before funding was discontinued altogether. In addition to the two road sections that had been completed prior to this decision, many other road projects had completed rehabilitation designs but had not been rehabilitated. However, the Armenian government used financing from the World Bank to rehabilitate many road sections that were included in the RRRP plans before they were discontinued, using and updating many road project designs developed by MCA. The Armenian government initially received a $25 million World Bank loan to rehabilitate road sections, which complemented the government's own investment of $16 million to rehabilitate approximately 50 km of the LRN. In August 2009, the World Bank approved additional financing of $36.6 million in loans to rehabilitate another 146 km of LRN, and in 2010, the World Bank approved another $40 million in low-interest loans. This brought the Bank's total lending for road rehabilitation of over $100 million to rehabilitate a total of 430 km of rural roads, most of which came from the original 943 km planned in the Compact.[6] As illustrated in Table I.1, the final MCA-funded portion of the RRRP was $8.4 million, compared to the final World Bank–funded allocation of over $100 million and the government of Armenia's contribution of $16 million.

[6] According to a World Bank source, a small portion of these 430 km of rehabilitated roads were outside the original 943 km identified in the original MCC-funded RRRP, but the exact degree of overlap was not known.

Table I.1. Summary of RRRP investments and targets in Armenia

	Compact	Rescoped MCA Project (2008)	Final MCA Project (2009)	Armenian Government/World Bank Project (2008–2013)
Funding (US$, millions)	$67	$67	$8.4 in MCA funds	$101.6 in World Bank funds; $16 in RA funds
Rehabilitated road links	Up to 943 km	297	24.4 km	430 km
Beneficiary communities	265	84	12	Not tracked by the World Bank
Beneficiary residents	360,000	142,905	6,356	Not tracked by the World Bank

Source: MCC-Armenia Compact 2007, Socioscope 2010: Rural Roads Rehabilitation Project Qualitative Process Analysis Final Report, and World Bank 2013.

RA = Republic of Armenia

The MCA and the Armenian government/World Bank road rehabilitation projects took place in the context of the world financial crisis, which had a dramatic impact on the Republic of Armenia. In 2009, GDP fell by about 16 percent. In particular, the construction sector experienced large losses, with employment declining by 40 percent. As a result, many migrant workers in Russia lost their jobs and returned to rural Armenian communities, further contributing to oversaturated local labor markets (World Bank 2010). By mid-2009, the poverty rate had increased to 28 percent, up from 24 percent in 2008 (World Bank 2010).

In this context, the World Bank–financed road rehabilitation program had a stronger focus on short-term job creation than the MCA-funded project. The program logic for the World Bank–financed rehabilitation of rural roads was mostly similar to MCA's program logic, but because of poor economic conditions at the time, one important difference was the World Bank's explicit objective of generating temporary employment related to road construction contracts as an immediate outcome of the rehabilitation work. The World Bank estimated that 36,650 person-months of employment would be created over the project's lifetime from 2009 to late 2011 (World Bank 2013), one to two years after they would have been rehabilitated under the Compact. Besides those individuals who would gain employment through the rehabilitation project, the Bank identified residents living within 2 km of the project roads as direct beneficiaries of the project, and people (living over 2 km from the project roads) who could feasibly use the roads to access social and economic services as indirect beneficiaries of the project.

The original scope of the evaluation—estimating impacts of the Compact-funded RRRP on beneficiary communities—was no longer possible once MCC discontinued funding, but the central question of whether road rehabilitation had impacts could still be examined. Most importantly, the data still covered many roads that were rehabilitated as well as a comparison group of roads that were not rehabilitated. The key differences from the original evaluation design are (1) the road projects were originally designed by MCA but financed by the World Bank and the Republic of Armenia (instead of funded by MCA-Armenia); (2) the follow-up period was shorter, only one year in some cases, so there is less emphasis on medium- and long-term effects; and (3) the evaluation methodology is different, as we discuss further in subsequent

sections. Because the fundamental question was unchanged, and because MCC had already committed to fund the data that would be used, MCC decided to proceed with funding the evaluation of the RRRP.

The report proceeds as follows: In the remainder of this chapter, we provide an overview of existing literature on rural road rehabilitation and the study's research questions, data, and evaluation design. In Chapter II, we provide a summary of RRRP implementation and the sample of households in the evaluation. In Chapter III, we present the impacts of the RRRP, and we summarize these impacts in Chapter IV. For ease of exposition, we hereafter use "RRRP" to refer to the set of roads projects as they were actually implemented—that is, the set of rehabilitation projects whose designs were funded by MCA-Armenia but whose construction was financed by the World Bank.

B. Prior research on road rehabilitation projects

Many roads that connect rural communities to larger cities and to each other in the developing world are in poor condition, constraining rural households from engaging in economic opportunities that would improve their well-being. The time and vehicle damage incurred in traveling along such roads make transport prohibitively costly, and at certain times of the year, such as during snowy or rainy seasons, some roads become impassable. Agriculture is a major economic activity in many rural communities in developing countries, and roads are hypothesized to be especially important for farming households. Roads that are in poor repair may restrict farming households from markets where they can buy agricultural inputs or sell their products, leading to potentially higher input costs and lower prices for crops sold. Repairing roads is intended to increase the profitability of farms and also expand the nonagricultural employment opportunities for rural households, increasing household income and, in turn, increasing household consumption and decreasing poverty rates. Repairing rural roads is also thought to improve access to social infrastructure such as medical services and schools. For these reasons, foreign aid agencies have made substantial investments to rehabilitate rural roads in developing countries (BenYishay and Tunstall 2010; Estache 2010; van de Walle 2009; Dercon et al. 2008; Escobal and Ponce 2002; Gannon and Liu 1997).

A growing literature attempts to estimate causal impacts of investments in rehabilitating rural roads in many countries. For example, van de Walle and Mu (2007) examine rural roads in Vietnam; Gibson and Rozelle (2003) examine rural roads in Papua New Guinea; Escobal and Ponce (2005) examine rural roads in Peru; Lokshin and Yemtsov (2005) examine rural roads in Georgia; Dercon et al. (2007) examine rural roads in Ethiopia; and Khandker et al. (2008) examine rural roads in Bangladesh. These studies have generally found that rehabilitating rural roads improves households' material well-being, though the studies vary somewhat regarding which channels drive the impacts. Escobal and Ponce (2005), for example, find that nonagricultural income increases but agricultural profits do not, and the effect on income does not lead to increases in consumption, possibly because households view the road improvements as transitory. However, many of the other studies find that road improvements lead to gains in agricultural income (increases in the prices received and volume of crops produced by farmers) as well as wages received for nonagricultural employment. Most of these studies focused on longer-term impacts of four years or more, or in some cases, estimated impacts for a set of roads with follow-up periods ranging from a few months to several years. Of the aforementioned

studies, only van de Walle and Mu (2007) consider impacts over a time horizon comparable to the present study (albeit in a much different context). Although their analysis focuses on impacts approximately four years after rehabilitation, they also examine impacts at a point in time where roads had been rehabilitated for an average of just over two years. They find that very few of the intermediate or longer-term impacts that were observed after four years were evident after two years.

The feature distinguishing these causal impact evaluations from less rigorous attempts is that they are better able to establish the counterfactual—or what would have happened in the affected areas if rural roads had not been rehabilitated. In most cases, this is done using a non-experimental estimation strategy in which the researchers have data for the "treatment" group—for example, communities and households served by the rehabilitated roads—measured before and after rehabilitation, as well as corresponding data for another "comparison" group served by roads that were not rehabilitated. The comparison group is crucial for providing a measure of the treatment group's counterfactual. The general approach is to use some method, such as statistical matching, instrumental variables, or predetermined observable criteria, to construct a comparison group that is similar to the treatment group (or sometimes a subset thereof) along observable dimensions.

Baseline data are crucial for determining how similar the treatment and comparison groups actually are, at least for observable dimensions. If the two groups are dissimilar, additional statistical matching or statistical modeling are employed to adjust for differences in initial conditions between the two groups. However, underlying these non-experimental methods is the assumption that there are not unobserved factors that are correlated both with whether a given community is in the treatment group and the outcome measures of interest. Van de Walle (2009) points to several potentially culpable factors that are unlikely to be observed but could bias impact estimates in non-experimental studies of roads, including political power, local leadership, social capital, and social empowerment. Fortson et al. (2013a) test several popular non-experimental methodologies and conclude that with rich baseline data, non-experimental estimators can greatly reduce bias in impact estimates, but some bias likely remains. The context in the Fortson et al. study is quite different—they consider the impacts of U.S. charter schools on student achievement—but the baseline data in that setting are likely to be more highly predictive of subsequent outcomes and less susceptible to omitted variable bias than are the household survey data usually available in evaluations of rural road rehabilitation projects.

The preferred impact evaluation approach for evaluating the impacts of rehabilitating rural roads would be a randomized-controlled trial (RCT) in which rural roads under consideration for rehabilitation are randomly assigned to a treatment group that will be rehabilitated and a control group that will not. The key distinction between a well-implemented RCT and a non-experimental evaluation strategy is that, because assignment to the treatment group is random in an RCT, treatment status is not correlated with unobserved factors that bias the impact estimates (Banerjee and Duflo 2009; Imbens 2010). To our knowledge, no researchers have conducted an RCT of a rural road rehabilitation project, though in an important contribution to the broader literature on road projects, Gonzalez-Navarro and Quintana-Domeque (2012) examine the impacts of a road-paving program in an underdeveloped area of a Mexican city using an RCT. RCTs have not been used to evaluate rural road rehabilitation projects mainly because foreign aid agencies understandably choose to fund the roads that are projected to provide the greatest

net benefits (BenYishay and Tunstall 2010) rather than randomly select projects. Roads might be selected based on estimated economic rates of return, number of beneficiaries who would be served, or perceived demand for the project. However, considering the frequency and scale of road rehabilitation programs, it is our hope that an aid agency will consider using an RCT to rigorously evaluate a future road rehabilitation program, perhaps randomizing among a larger set of road projects that would qualify for funding but cannot all be served within a set budget.

Despite the caveats of the aforementioned non-experimental studies, they are still the best evidence we have about, and provide valuable insights into, the possible impacts of rural road rehabilitation. The present study also relies on a non-experimental evaluation approach that is subject to the same concerns itemized above. The main contribution of the present study is to provide further evidence on the short-term impacts of rehabilitating rural roads based on a project implemented in a new context and with a broader set of outcomes than has been examined in most of the prior studies.

C. Evaluation questions, data, and estimation method

The key research questions guiding our impact evaluation of the RRRP evaluation are as follows:

- ***Did rehabilitating roads affect the quality of roads?*** We would expect road quality to be affected immediately, and reflected in residents' assessments of the quality of regional roads.
- ***Did rehabilitating roads improve access to markets and social infrastructure?*** Several indicators allow us to measure access, including residents' accounts of market access; the time it takes to travel to hospitals, markets, and schools; and the availability of transportation. Access is hypothesized to be impacted in the short term.
- ***Did rehabilitating roads improve income from employment?*** Short-term construction employment was a key objective of the World Bank and is hypothesized to be immediately observable. Improved roads are also hypothesized to increase employment income in the longer term by increasing access to job opportunities, but this effect is unlikely to have fully materialized in our follow-up time frame.
- ***Did rehabilitating roads affect agricultural productivity and profits, and if so, by how much?*** We measure potential effects of the RRRP on the volume and variety of agricultural and food production, as well as income from agricultural and food sales. Impacts on these outcomes are not expected to fully materialize until a few years after roads are rehabilitated, and with the short follow-up period of the present report, we would not expect large impacts but might observe early improvements.
- ***Did rehabilitating roads improve household well-being for communities served by these roads, especially income and poverty?*** Likewise, these outcomes are only expected to fully manifest in the medium and long term, but we could potentially observe early impacts.

These questions are particularly relevant to the Armenian context because the transportation sector plays a critical role in their economy. Goods and services related to transportation and communications represent an average of 6.7 percent of GDP over the last five years, and these

sectors have employed about 5.6 percent of the total workforce in 2011 (World Bank 2012). In addition, the agricultural sector represents about 20 percent of the country's GDP and involves about 335,000 households (World Bank 2012). Trade from rural areas is less than optimal due to restricted connectivity to markets as a result of the poor condition of roads. In particular, poor rural roads have generated substantial crop losses in numerous communities, as farmers are unable to transport their crops to market in a timely manner. Theoretically, mitigating the structural constraint of poor rural transportation infrastructure could improve and promote agricultural trade, thus stimulating economic growth and local employment. The RRRP was designed to meet these long-term objectives.

Data. Data availability had a large bearing on the estimation method for this impact evaluation. The data for the RRRP impact evaluation come from the Integrated Living Conditions Survey (ILCS). The ILCS is an annual, nationally representative household survey fielded by the National Statistical Service of Armenia. The ILCS covers a broad range of topical domains, including demographics, employment, income, agriculture, and public services. The core sample of the ILCS includes 768 enumeration areas, each containing 8 households for a total sample of approximately 6,100 households. The survey is implemented year-round; all 8 households from a given enumeration area are interviewed for the month in which the enumeration area is selected. It is important to note that the ILCS is cross-sectional, with a new cross-section of respondent households (and enumeration areas) drawn each year. This is in contrast to longitudinal household surveys, which sample the same households from year to year.[7]

The ILCS featured a larger sample of communities and a longer survey questionnaire from 2007 to 2011 to facilitate the originally-designed RRRP evaluation. During these years, the ILCS oversampled communities that would be directly connected to the LRN through rehabilitation of MCA's initial set of eligible roads. This oversample comprised 216 enumeration areas in rural communities, or approximately 1,700 additional households. Additional communities served by rehabilitation-eligible roads were selected into the core sample by chance, and a total of approximately 2,200 households served by the original eligible project roads are in the sample each year.

The full ILCS sample includes communities served by 82 of the 85 original eligible roads. In addition, the standard ILCS questionnaire was expanded during the evaluation period to include questions about intermediate outcomes related to the RRRP as well as an expanded set of final outcomes, as described below. This expanded questionnaire—combined with the oversample of communities connected to rehabilitation-eligible roads—better tailored the ILCS to the evaluation's research questions and sample size requirements.

[7] Because the same households were not surveyed from year to year, we are not able to control for baseline household-level characteristics that could be correlated with the evaluation's outcomes of interest—namely each household's total income and consumption at baseline. However, we can control for baseline community-level characteristics—namely the average income and consumption of all sampled households in each community—as described on page 14.

ILCS data from 2007 and 2008 precede program implementation and can thus be used as baseline data.[8] ILCS data from 2011 were collected at approximately one year after construction in 2010 and two years after construction in 2009. As such, these 2011 data serve as follow-up data for communities served by the RRRP in 2009 and 2010.

In addition to ILCS data, the evaluation will use original ex ante ERRs calculated for all 85 road links that were assessed by stakeholders for inclusion in the RRRP. We use these ERRs as control variables in the regression described below as well as additional sensitivity tests. Additional ERRs were calculated at several points in time as part of the design, rescoping, and implementation of the RRRP and are available for a subset of road links. We do not use these ERRs, however, because they are not available for all road links in the evaluation and few of the comparison road links.

Evaluation sample. Our sample includes communities near road links that were improved by the RRRP in 2009 and 2010, as well as similar communities near road links that were eligible for the RRRP but not improved in the same time period. To minimize sampling variability due to cross-community differences, we further restricted the communities included in the analysis to those that have 2011 data and at least one year of preintervention data: either 2007 or 2008, or both years. Of the 82 road links in the full ILCS sample, there are 60 road links that meet these criteria: 30 treatment links and 30 comparison links. However, a small number of road links are connected to each other; we treat each of those pairs as a single road that gets double the weight of each other individual road (hereafter, we refer to these combined roads as single units for simplicity of exposition). All told, our analysis sample includes 27 treatment road links and 28 comparison road links, and over 50 communities and 2,300 households in each of the treatment and comparison groups (Table I.2).[9,10] Of the 27 treatment links, 17 underwent rehabilitation in 2009 and 10 underwent rehabilitation in 2010. In contrast, none of the 28 comparison links underwent road rehabilitation during the study period. Unfortunately, ILCS data were unavailable for households and communities affected by MCA-funded rehabilitation of the pilot roads. As a result, all 27 treatment links in the evaluation sample were funded with World Bank loans to Armenia.

[8] We can use both years of preintervention (baseline) data on key outcomes—rather than just one year—to improve the precision of impact estimates.

[9] Of the 55 communities in the treatment group, 39 had both years of baseline data and 16 had only one year of baseline data. Of the 52 communities in the comparison group, 34 had both years of baseline data and 18 had only one year of baseline data.

[10] Three project roads were excluded to preserve geographic balance in the ILCS oversample. In addition, communities served by a similar foundation-funded project are excluded, as are communities that received rehabilitation assistance in 2011, the year of the follow-up survey.

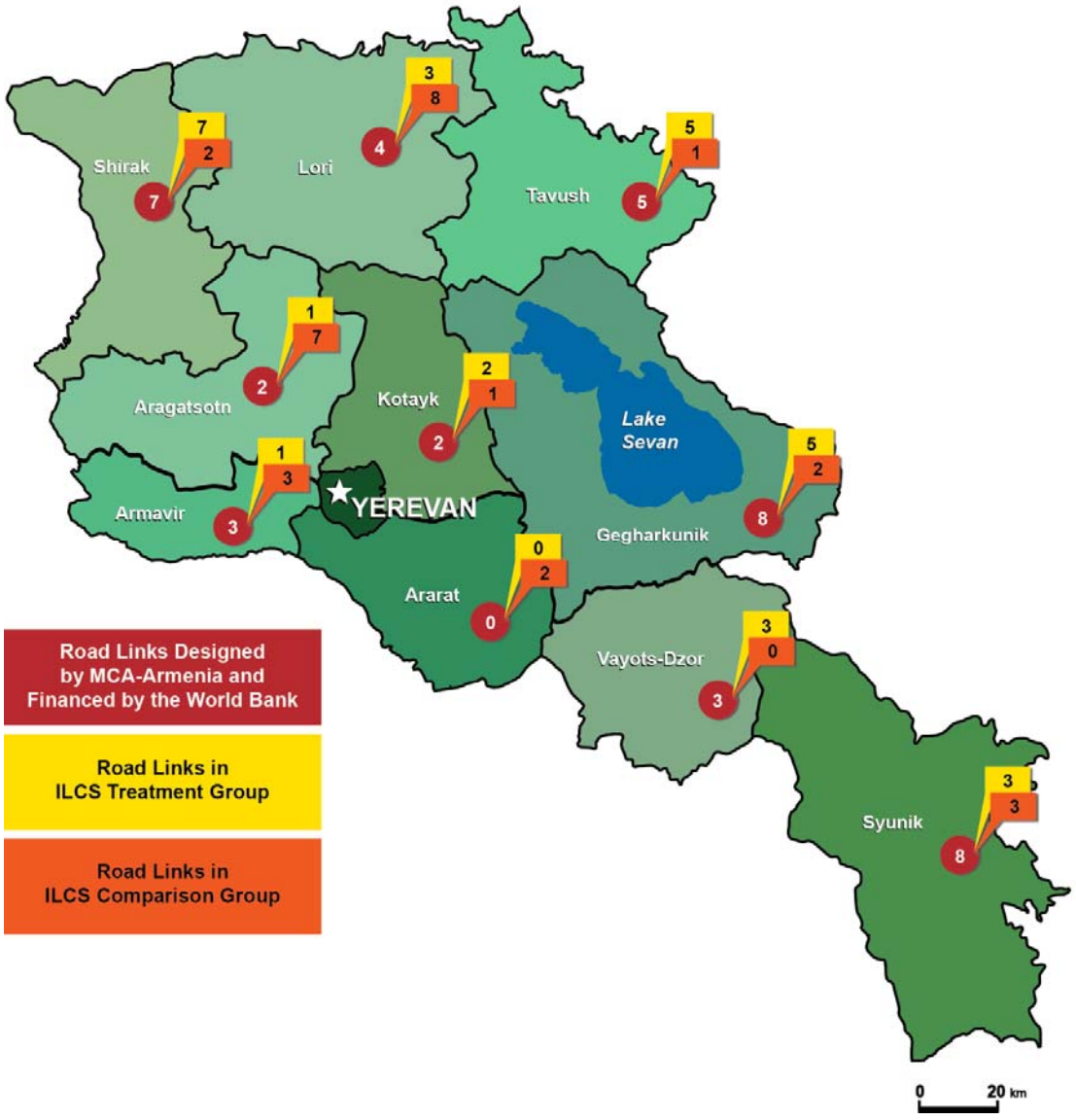
Table I.2. Links, communities, and households in the evaluation sample

	Comparison Group	Treatment Group	Treatment Group (construction in 2009)	Treatment Group (construction in 2010)
Road Links	28	27	17	10
Communities	52	55	29	26
Households	2,288	2,560	1,152	1,408

As illustrated in Figure I.3, road projects and the ILCS sample coverage span most of Armenia. All marzes,[11] with the exception of Ararat, had at least one road link designed by MCA-Armenia and financed by the World Bank. Similarly, all marzes have at least one treatment or comparison road link in the evaluation, and all but two marzes have at least one road link in the treatment group and one in the comparison group.

[11] Armenia is divided into 11 administrative divisions. Of these, 10 are marzes, or provinces, and the country's capital, Yerevan, is granted special administrative status as the country's capital.

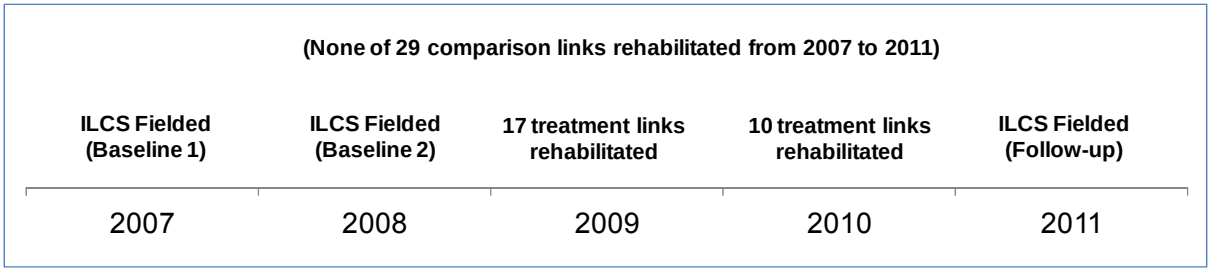
Figure I.3. Distribution of treatment and comparison road links



As mentioned above, we use 2007 and 2008 ILCS data as baseline data for this analysis, and 2011 ILCS data as follow-up data (see Figure I.4 for an overview, and Figure II.3 in the next chapter for details). Baseline data for treatment group communities were collected between one and three years before roads were rehabilitated—three years in the case of treatment links rehabilitated in 2010 for which 2007 ILCS data are available. Similarly, follow-up data for treatment communities were collected approximately one to two years following road rehabilitation for most households, depending on whether road links were improved in 2009 versus 2010 and when within the year the household was interviewed. For some households interviewed in early 2011, the time between the survey and rehabilitation could be as little as 6 or 7 months if rehabilitation was completed in 2010, but for most it was at least one year. As a

result, this evaluation will generate an estimate of the impact of the RRRP approximately one to two years following construction of most road links.

Figure I.4. Summary timeline of implementation and survey dates



Estimating the counterfactual. A crucial objective of any impact evaluation is to assess not only how key outcomes for affected beneficiaries changed after the intervention, but also to assess the counterfactual: how outcomes *would have changed* even in the absence of the intervention. Randomly assigning the intervention is the ideal way to identify a credible counterfactual. However, as is frequently the case for large-scale infrastructure projects, it was not feasible to randomly select the roads that would be rehabilitated in the MCA-Armenia RRRP from among the eligible. Instead, MCA-Armenia planned to fund projects based on estimated economic rates of return. The ERR is calculated from several inputs, including the vehicular traffic, vehicle operating costs, and the cost of the project. An ERR was estimated for each of the road links under consideration for rehabilitation, and a road's ERR had to meet or exceed 12.5 percent for the road to be funded.

We initially planned to exploit the systematic selection of roads for rehabilitation to implement another rigorous evaluation design, a regression discontinuity approach, but this ultimately was not possible. In a *regression discontinuity* design, treatment is systematically assigned based on a known "score," which in this context is the ERR; conceptually, impacts are measured by comparing outcomes for those above and below the score selection threshold. Because the exact selection criteria are known and can be accounted for in the statistical analysis, regression discontinuity is considered a strong, credible alternative when random assignment is not possible. In this context, the research design would have essentially compared roads that just passed the ERR threshold to those that fell just short of it. The roads on either side of this cutoff have very similar ERRs, yet roads above the threshold would be rehabilitated, and roads below would not be. For roads near the threshold, ex ante differences between the two groups are minimal, and therefore outcome differences for roads on either side of the threshold would reflect program impacts. The data were set up to facilitate this evaluation design.

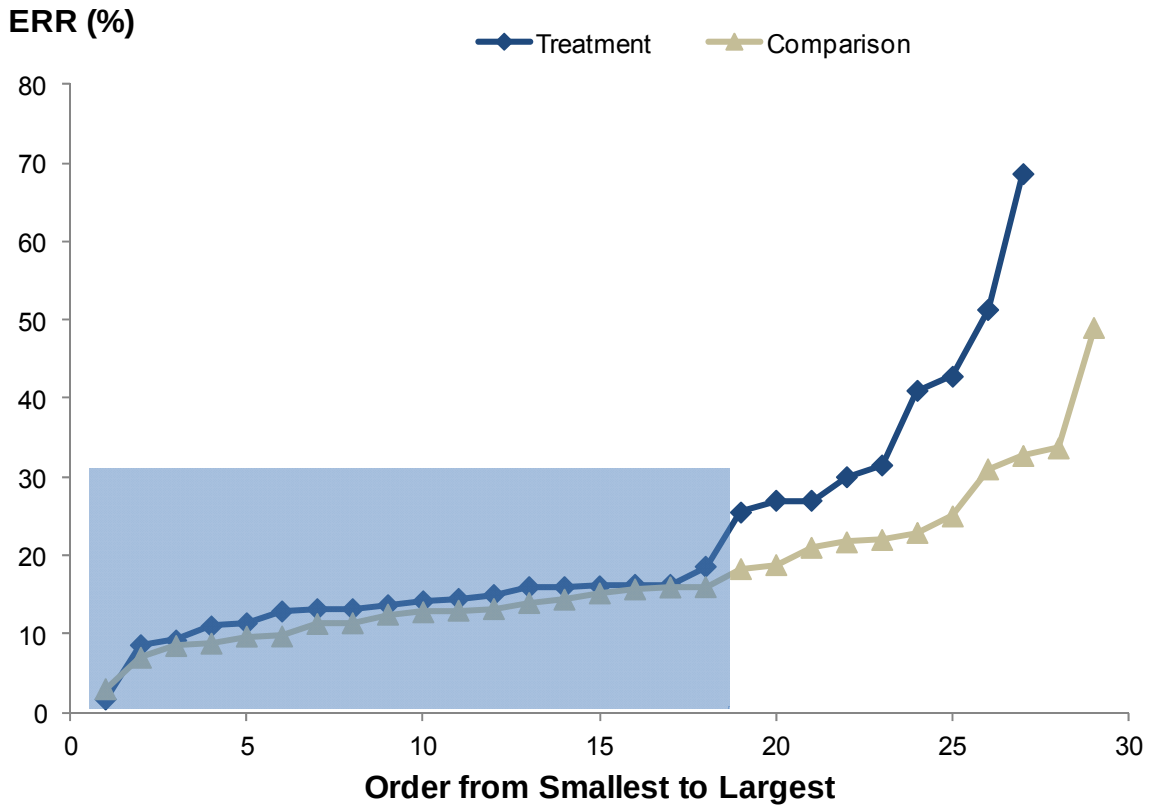
After the MCA road rehabilitation program was put on indefinite hold, the World Bank agreed to finance rehabilitation of some of the road links that had been included in the RRRP plans. (The Bank also concurrently funded other road rehabilitation projects that were not considered under the RRRP.) The World Bank based the rehabilitation efforts on project designs developed by MCA, with some of the designs updated due to changes in conditions after the designs were developed about two years prior. Implementation of these projects began in late 2009 and, in conjunction with subsequent funding, continues into 2013, though the present report

focuses on projects completed by 2010.[12] Our evaluation focuses on road links that were formerly in the RRRP but were ultimate financed by the World Bank. Like MCA-Armenia, the World Bank requires that the projects it funds be economically justified with sufficiently high ERRs. However, in selecting which of MCA-Armenia's original project designs would be funded by the World Bank, they used ERRs as well as additional criteria, particularly regarding the number and proportion of residents near roads that were involved in commercial and agricultural practices. The result was that many projects with low ex ante ERRs were selected and many with relatively high ex ante ERRs were not selected. Consequently, the regression discontinuity evaluation design we developed was no longer a viable approach to measuring the impact of road rehabilitation.

Given that a regression discontinuity design was not feasible, we use a simpler comparison group evaluation design. The treatment group in our analysis comprises the 27 road links that were in MCA-Armenia's original RRRP plans, were ultimately rehabilitated with financing from the World Bank, were completed by 2010, and were covered by the ILCS from 2007/2008 to 2011. The comparison group in our analysis comprises the 28 road links that were in MCA-Armenia's original RRRP plans, were not ultimately rehabilitated with financing from the World Bank as of 2011, and were covered by the ILCS from 2007/2008 to 2011. Our data include outcomes measured before any roads were rehabilitated and after they were completed for both the treatment and the comparison group road links.

[12] A subset of the analysis, examining impacts on short-term employment, includes road links rehabilitated in 2011 as well.

Figure I.5. Initial economic rates of return of treatment and comparison road links



As illustrated in Figure I.5, the highlighted 18 treatment and 18 comparison links with the smallest ERRs track each other closely. However, the highest ERRs among treatment links are notably higher than their counterparts among the comparison group, by about 10 percentage points on average. Our regression models are set up to capture the potential importance of the ERRs in determining treatment status, as this was expected to be an important predictor of selection, though the relationship between ERRs and treatment status is weaker than we expected. Although ERRs were not as important to the model as we expected, they improved the model's fit compared with a simpler model that excludes ERRs. We also conducted sensitivity analyses in which we limit the analysis to the highlighted road links, and in which we exclude ERRs from the model altogether; these are reported in Appendix A.

Although the regression specification described in more detail below is similarly structured to a regression discontinuity design, because ERRs are only weakly predictive of treatment status, the evaluation approach is best thought of as a comparison group or "difference-in-differences" design. As described previously, an important concern with a comparison group design is that, because the treatment and comparison roads are not randomly selected, there may be observed or unobserved differences on some dimensions that affect the key outcomes. More specifically, the reasons a given road was selected for rehabilitation may also contribute to differences in average outcomes for the two sets of road links. To the extent that the comparison group differs from the treatment group along dimensions important for outcomes, we will be unable to distinguish actual program impacts from underlying differences in the two groups

unless we can credibly identify and control for these preexisting differences. For example, households in the treatment group could have higher annual earnings than households in the comparison group. We must control for these higher earnings to differentiate them from any additional household income resulting from road rehabilitation.

Regression model. In this section, we describe the empirical model that we used to estimate impacts of the implemented road rehabilitation projects, given that we have baseline data from 2007 and 2008 and follow-up data from 2011 for all treatment and comparison communities in the sample. As mentioned above, we used a difference-in-differences framework to analyze RRRP impacts. The central idea behind difference-in-differences estimation is to estimate how outcomes changed for communities served by treatment roads before and after road rehabilitation (the first difference) and compare this to how outcomes changed for communities served by comparison roads over the same time period (the second difference). The difference between these two differences is the estimated impact that can be attributed to the program. Impacts can be estimated mathematically using the following regression model:

$$(1) \quad y_{irt} = \beta' x_{ir} + \theta' R_t + \lambda T_r + \gamma R_{t=2011} T_r + \delta' R_t \times ERR_r + \eta_{rt} + \varepsilon_{irt}$$

where y_{irt} is the outcome of interest for household i served by road link r at time t; x_{ir} is a vector of time-invariant characteristics of household i served by road link r as well as road link–specific characteristics; T_r is an indicator equal to 1 if road link r is in the treatment group and 0 if it is in the comparison group; R_t is a vector of binary variables for each round of data included in the analysis: 2007, 2008, and 2011; ERR is the economic rate of return associated with each road link at baseline; η_{rt} is a road-specific error term; ε_{irt} is a random error term for household i served by road link r observed at time t; and β, θ, λ, γ, and δ are parameters to be estimated. The vector of baseline characteristics x_{ir} includes household size and composition and characteristics of the household head—namely, education level, gender, and age. The model also controls for geographic location, namely the marz in which surveyed households were located. The vector R_t accounts for countrywide trends that affect all households in a particular year. The estimate of the coefficient λ is an estimate of how different outcomes were in the treatment and comparison communities prior to rehabilitation.

In this formulation, the estimate of the parameter γ, the coefficient on the interaction of treatment status and the year 2011 indicator variable, is the difference-in-differences impact estimate and represents the difference in outcomes for treatment and comparison communities in 2011, the first year after projects are complete in all of the beneficiary communities, relative to how different they were prior to rehabilitation. We note that our regression model differs slightly from a conventional difference-in-differences framework in that the pre-intervention years vary by community. For some communities we only have pre-intervention data for 2007, for others only 2008, and still others have data for both years. In our framework, the pre-intervention period is 2007 or 2008 if a given community only has data in one of the years, and it is an average of the two if the community was surveyed in both years.

Equation (1) was estimated using ordinary least squares.[13] We use sampling weights constructed by the National Statistical Service that accounted for the number of households sampled and the population of communities surveyed in each road link. The total weight across households for each road link is the same for all links. Household weights were rescaled to give each road link equal weight, and to give each of the road link pairs that were combined double the weight of individual road links, as described previously. Because all households served by a particular road link are assigned to treatment status collectively, we needed to account for the fact that those households may have had correlated outcomes, represented by the road link–specific error term in Equation (1). As an example, a particular community might have abnormally good or bad weather, or could experience other economic "shocks" that are unrelated to the road project but nonetheless affect all communities served by that road. Road-level correlations were accounted for using Huber-White standard errors, and these errors are clustered by road link and year.

Under the ILCS sampling design that was implemented for this study, the treatment group only includes communities that were originally thought to be served directly by the project. However, other nearby communities may benefit from rehabilitation as well, because impacts are likely largest in communities directly served by project roads. Hence, the estimated impacts will be interpreted as the average impact on households in communities that are directly served by project roads, as opposed to the average impact for all households affected by the project roads.

Impacts on subgroups. For many of the outcome measures, it is conceivable that the effects of the road projects will vary by observable household characteristics.[14] Estimating differential impacts on female-headed households is of particular interest to MCC, so we examine whether the interventions' effects differ for the subgroup of female- and male-headed household. It is straightforward to embed subgroup estimates into the framework of our previous specification. To do so, we include an interaction term that distinguishes treatment group members in subgroup S from those who are not in the subgroup:

$$(2)\quad y_{irt} = b\phi x_{ir} + q\phi R_t + l_{S=1}T_r'(S_{ir}=1) + g_{S=1}R_{t=2011}T_r'(S_{ir}=1)$$

$$+\lambda_{S=0}T_r \times (S_{ir}=0) + \gamma_{S=0}R_{t=2011}T_r \times (S_{ir}=0) + \delta' R_t \times ERR_r + \eta_{rt} + \varepsilon_{irt}$$

[13] We use OLS to estimate impacts on binary outcomes as well. When applied to binary outcomes, OLS has two theoretical problems. The first potential problem is that predicted probabilities may be less than 0 or greater than 1. The second problem is that the error terms in the model will violate distributional assumptions, in which case statistical inference could be incorrect. To overcome these problems, researchers often use probit or logit models to estimate impacts when the outcome measure is a binary variable. However, probit and logit models have problems of their own, and the theoretical problems of an OLS model with binary outcomes are rarely problematic in practice. (See Appendix A of Fortson et al. (2013b) for a more detailed discussion.) We tested several binary outcomes using OLS and logit models to confirm that the impacts and significance levels were almost identical with either approach.

[14] In the original evaluation design, we also suggested subgroup analysis for subgroups defined by road link characteristics (such as region), not just by household. However, because of the smaller number of road links that were rehabilitated with World Bank funding, it was not feasible to estimate impacts separately for subgroups of road links.

In Equation (2), the estimate of γ_S=1 is the estimated impact for members of subgroup *S,* and likewise, γ_S=0 is the estimated impact for households that are not in subgroup *S.* We can test whether the impacts differ for members of that subgroup compared to households not in that subgroup by statistically testing whether γ_S=1 and γ_S=0 are equal.

Ideally, we would also look at differential impacts for other subgroups, particularly the possibility of differential impacts for subgroups defined by preintervention household income. For example, poorer households might benefit more from road rehabilitation because they are most in need of better economic opportunities, or they might benefit less well because they lack the appropriate resources (such as vehicles) to take full advantage of the rehabilitated roads. Unfortunately, the cross-sectional nature of the data precludes estimating impacts for subgroups defined by time-varying characteristics because, for the households surveyed in 2011, we cannot classify them according to their preintervention income levels.

Sensitivity analyses. We explored six alternative regression specifications:

- Alternative 1: A simple difference-in-differences model, with no controls for any other household or road link variables. This model determines the extent to which controls for the interaction between ERR and year are altering estimates.
- Alternative 2: A model similar to Equation (1) but excluding the ERR by year interactions. This model examines the extent to which controls for the interaction between ERR and year are altering estimates.
- Alternative 3: The same model used in our main specification but restricted to the 18 treatment roads and 18 comparison roads with the smallest ERRs, which, as shown in Figure I.5, are similar for both groups. This model examines if keeping only those roads in either the treatment or comparison group with an ERR similar to that of a road in the other group results in substantially different impact estimates.
- Alternative 4: Inverse probability weighting on propensity scores estimated from road link–level baseline averages of key outcome measures and ex ante ERRs. This model examines if using propensity scores to reweight the sample as if there were a similar distribution of baseline characteristics between the treatment and comparison groups results in substantially different impact estimates.
- Alternative 5: A similar propensity score approach that also trims the treatment and comparison roads to the subset with similar estimated propensity scores. This model is similar to Alternative 4, but also explores whether restricting the analysis to the treatment and comparison road links that are most similar to each other affects the estimates.
- Alternative 6: Restrict the analysis to roads that were rehabilitated in 2009, so that there were at least two years between road rehabilitation and the follow-up survey year.There were few substantive differences in the impacts estimated using these alternative models. Appendix A describes the sensitivity analyses in more detail and compares the estimated impacts of the alternative specifications to the main model.

Design limitations. A primary limitation of the difference-in-differences estimation approach described above is that there may be systematic differences between treatment and comparison communities—besides access to rehabilitated roads—that could affect key outcomes in the evaluation. To the extent that the comparison group differs from the treatment group along dimensions important for outcomes, we are unable to distinguish actual program impacts from underlying differences in the two groups unless we can credibly identify and control for these preexisting differences. As van de Walle (2009) notes, several factors that we do not observe in the data but could bias impact estimates in comparison studies such as ours include political power, local leadership, social capital, and social empowerment. Neither the present study nor past evaluations of rural road rehabilitation projects conclusively overcome this concern, and it should be borne in mind as a potential source of bias in our findings. Additionally, the estimation strategy accounts for preexisting differences in baseline outcomes, but not preexisting differences in the *trends* of those outcomes. This is an especially important assumption because, as will be shown in Chapter II, treatment households had somewhat lower incomes, on average, than did the comparison group at baseline. We unfortunately do not have sufficient pre-rehabilitation data to test whether the treatment and comparison communities had similar growth rates for the key outcomes, such as household income and poverty.

D. Outcomes and statistical power

Although most of the outcomes of primary interest to MCA and MCC are longer-term outcomes, such as economic improvements (including household income), these outcomes are not immediately observable in the relatively short follow-up period available for this evaluation. Consequently, we focus on intermediate outcomes through which the road rehabilitation projects are intended to improve household income; these outcomes are more likely to change over the limited study period. We would expect a longer-term impact on households' income only if we observe that a substantial proportion of the targeted communities are actually experiencing improvements in their roads. Examining intermediate outcomes also establishes the counterfactual—how the quality of and access to roads would have changed even in the absence of the road projects.

Although we would not expect impacts on longer-term outcomes to have fully materialized in the time frame of the evaluation, these are central to the ultimate goals of the RRRP, and hence, household income and its components are still included in our analysis to assess if there is evidence of early impacts. The primary domains we focus on are agricultural revenue, costs, and employment income by household members, as improvements in these outcomes were central to MCA's conceptual framework for the Compact as well as to the World Bank's objectives. Rehabilitating roads might also increase nonagricultural income, especially employment opportunities; therefore, we also estimated impacts on key sources of nonagricultural income. Table I.3 summarizes the key final outcomes that we examine using the ILCS data, as well as a short description of how they are measured.

Table I.3. Key outcome measures of the Roads Evaluation

Outcomes	Description	Link to Program Logic
Perceived quality of roads	Quality of roads both within the community and roads connecting the community to other communities; quality of local public transportation	Measure of immediate outcome of improved road quality
Availability of transportation	Modes of transportation are available and commonly used, especially to access key community services	Measures of short-term outcome of increased vehicular activity
Access to social infrastructure and markets (distance and time to school, clinics, etc.)	Distance and time from key services, including health facilities, schools, community centers, and markets	Measure of short-term outcome
Utilization of local roads	Whether (and how often) road transportation is typically used to purchase agricultural supplies, to sell agricultural produce, to access employment outside the community, or for other purposes	
Agricultural investments	Total amount spent on inputs and machinery related to crops and animals	Measure of medium-term outcome of increased investment
Agricultural production and sales	Total amount of specific crops grown and sold, animals bought, owned, and sold; value of crop and animal sales	Measure of medium-term outcome of increased production
Income from employment	Income from any temporary or permanent employment	Component of long-term outcome of household income
Household income	Sum of agricultural profits, employment income, and remittances	Measures of long-term outcomes
Household consumption	Sum of all household expenses, including food expenses, house and car repair, education, and health expenses	Measure of long-term outcome
Household poverty	Whether the household is poor according to their annual consumption	Measure of long-term outcome

Precision of the impact estimates. A limitation of the impact analysis is that our estimates will not be very statistically precise because of features of the evaluation design. Relative to an unclustered sample design, the clustered nature of the intervention reduces statistical power. Instead of having households drawn independently from many different communities, they are drawn from just 56 road projects. Moreover, one of the key outcomes, annual household income, is highly variable, making it difficult to precisely estimate impacts even if our sample were large and unclustered. To illustrate statistical precision of the impacts, we calculated minimum detectable impacts—the smallest true impacts that can be reliably detected—based on the

standard errors from our analyses for three key outcomes.[15] We report minimum detectable impacts for three representative outcomes below, and minimum detectable impacts for a more comprehensive set of outcomes are reported in Appendix B.

First, we considered the minimum detectable impact on household income, which is the key long-term outcome for assessing the economic benefits of the RRRP. The minimum detectable impact is large, $1,092, which translates to 30 percent of the comparison group's average annual household income. The imprecision for the income impact was expected at the outset of the evaluation design and motivated a focus on other, less variable outcomes as well, particularly poverty and consumption.

Second, we considered the minimum detectable impact on poverty, which is the key long-term outcome that is central to MCC's mission to reduce poverty in developing countries. The minimum detectable impact on poverty is also large, 15 percentage points, which would mean poverty is reduced by more than half. Our ex ante calculations of the minimum detectable impact on poverty were high, but not as high: 11 percentage points. The difference is mainly attributable to clustering effects being more prominent than we had expected for this outcome. The limited statistical power for income and poverty (and other related outcomes) is especially important because we estimated impacts based on data collected one year after many of the programs were completed. Limited statistical power means that we will not be able to detect small impacts, as we might expect at the one-year follow-up (when full impacts have not been realized). We emphasize this limitation when we present the estimated impacts in Chapter III.

Last, because of the short follow-up period and limited statistical precision for longer-term impacts, we also considered the minimum detectable impact on perceived road quality, a key short-term outcome for which we would expect large impacts. Thus, this outcome can provide an indication that larger impacts on household well-being may develop in the future. The minimum detectable impact on respondents' assessments of their regional roads as good or better is 24 percentage points. Clustering is even more pronounced for this outcome, mainly because all households within a community have the same roads, so households will only vary in their perception of those roads. This is a large impact, but if the RRRP is successful, we would expect dramatic improvements on this dimension.

[15] The minimum detectable impact is calculated as the standard error for the impact estimate for the specified outcomes times 2.80, which is the factor associated with a two-tailed t-test with 80 percent power and a 5 percent significance level.

This page has been left blank for double-sided copying.

II. IMPLEMENTATION SUMMARY AND STUDY SAMPLE

In this chapter, we discuss the planning and implementation of MCA- and World Bank–funded road rehabilitation efforts, summarize construction outputs and costs, and present qualitative findings regarding implementation and project results.

A. Initial planning and road selection

The RRRP was designed through a consultative compact development process, in which more than 1,200 individuals participated and 230 written proposals were submitted for particular projects (BenYishay and Tunstall 2010). Improved physical infrastructure was highlighted as a priority area, so the Armenian government selected improvements in irrigation infrastructure and roads as the main areas for MCC support. In the road sector, the original proposal from the Armenian government to MCC included over 1,100 kilometers of rural roads that had been selected from the LRN through an initial screening that took into account population, estimated traffic, and distance to the capital (BenYishay and Tunstall 2010).

The extensive list of road links was pared down through due diligence as more information became available and more analyses were completed. To prioritize roads for rehabilitation, MCA-Armenia constructed ERR calculations for proposed road links. ERRs measured the monetized economic benefits of the proposed and actual rehabilitation of the rural road network, after accounting for project costs.[16] Stakeholders assumed that the primary benefit streams of road rehabilitation were (1) operating cost savings for vehicles and wear and tear on the roadbed, and (2) travel time savings, as improvements in road condition permit higher average driving speeds (MCC 2011). All of the roads that exceeded a hurdle rate of 12.5 percent[17] were included in the original program plans. A few exceptions were made for roads that provided the lifeline connection to communities included in the Irrigated Agriculture Project. At the end of this process, MCA-Armenia selected 943 km of 85 road links for the RRP, and this 943 km of roads was codified in the Compact.

Using this model, the ex ante ERR for the total RRRP was estimated at 26 percent over 30 years (MCC 2011). At over double the hurdle rate of 12.5 percent, this represented a highly viable investment. Aggregating costs and benefits for all roads selected for the RRRP, it was estimated that this project would result in reduction in transportation costs of $20 million a year beginning five years after material project benefits were realized (MCC 2011).

[16] These calculations were computed using the Highway Development and Management Model, (HDM-4), an economic model for evaluating ERRs of road projects based on traffic counts and road roughness measures. The model forecasts the expected changes in road quality (roughness), speed, and traffic. Those estimates are then used to calculate the expected savings in vehicle operating costs and travel time, which are monetized, aggregated, and compared to the cost of the initial investment plus ongoing maintenance and repair over the life of the project (typically 20 years).

[17] This hurdle rate of 12.5 percent was chosen because it represented Armenia's average real growth rate from 2005 to 2007.

B. Early implementation and program modifications

Before implementation began, rehabilitation targets were reduced by two-thirds due to the devaluation of the U.S. dollar against the Armenian dram and construction price escalation. As a result, 290 kilometers of rural roads were selected for rehabilitation under the RRRP project with the same total cost of $67.1 million. The packaging was based on new ERR calculations updated based on the results of feasibility studies and submitted by the feasibility/design consultant (Sweco International) and with guidance from MCA-Armenia RRRP team and MCC. The project was to be implemented with one pilot phase and three separate packages.[18]

Construction in the pilot phase of the project started in early 2008. The pilot included two sections of the H17 Armavir-Isahakyan-Gyumri road; these sections spanned 24.4 km and connected a number of villages in the marz of Aragatsotn with Gyumri, the country's second biggest city. A portion of these 24.4 kilometers came from the 943 kilometers of the LRN originally targeted by the RRRP.[19] The original rehabilitation designs for these pilot road sections were completed by the Lincy Foundation. The Foundation's scale-down of assistance in 2008 and 2009 led to road rehabilitation designs that were not executed. MCA used these existing rehabilitation designs, in part, to initiate operations quickly and concurrently with planning for full rollout of the RRRP. The MCA-Armenia RRRP team reviewed these designs and supervised the Armenian Road Directorate (ARD) as it paved and improved drainage systems for these pilot roads. In addition, several companies were involved in construction supervision, including the Swedish firm, Sweco International, and a consortium of a German firm, Kocks Consult GmbH, and an Armenian firm, Arosa Ltd. By June 2009, these two pilot sections were fully rehabilitated with MCA funding (see timeline in Table II.1).

In March 2008, the MCC Board decided to put a hold on MCC roads funding in Armenia, and then in June 2009, the MCC Board decided that MCC would not provide any further funding for road construction and rehabilitation. The primary reason for this decision was that a number of Armenia's political indicators had declined noticeably following the Armenian government's response to protests about the 2008 presidential election. These indicators form part of the eligibility criteria based on which MCC makes decision about funding country programs.

[18] See http://www.mca.am/en/mca_armenia/roads/ for more details on each of the three packages of the RRRP.

[19] According to MCA sources, a portion of these pilot road sections was included in the original 943 km of road links identified in the Compact, but another portion of pilot road sections was outside of these originally identified road links. However, the exact overlap between the 24.4 km and the 943 km is unknown.

Table II.1. Timeline of road construction and data collection

Activity	Date
Compact negotiations end with 943 km of rural roads for RRRP	2006
290 km of rural roads selected for rehabilitation following devaluation of the U.S. dollar	November 2007
First baseline round of ILCS	January–December 2007
MCA-funded rehabilitation of pilot roads begins	Early 2008
Second baseline round of ILCS	January–December 2008
World Bank approves initial $25M for RRRP	February 2009
MCC freezes funds for additional rehabilitation	March 2008–June 2009
24.4 km fully rehabilitated with MCA funds (pilot phase)	June 2009
World Bank–funded rehabilitation starts	June 2009
World Bank approves additional $36.6M for RRRP	August 2009
World Bank approves additional $40M for RRRP	July 2010
Follow-up round of ILCS	January–December 2011
A total of 446 km rehabilitated with World Bank Funds	December 2013

Source: World Bank 2014 and Socioscope 2010.

C. World Bank- and Armenian government–financed construction

In response to the 2008 world financial crisis, the Armenian government requested support from the World Bank to help weather the economic shocks. The Bank decided to utilize the "IDA Financial Crisis Response Fast-Track Facility" to support "shovel ready" investments which, in addition to Armenia's own initiatives, would create jobs and mitigate the negative impact of the financial crisis. The Lifeline Road Improvement Project was prepared in this context with the objective to upgrade selected sections of the LRN and create temporary employment in road construction. The board of directors approved the project on February 24, 2009. In this first round of funding, $25 million in World Bank loans were allocated for rehabilitation of the road sections included in the first of the three RRRP packages. This funding was designated for rehabilitating 100 km of the LRN and creating 7,600 person-months of temporary jobs. Also in 2009, the Republic of Armenia pledged $16 million to rehabilitate an additional 50 km of the original 943 km of the LRN.

The Armenian government/World Bank rehabilitation project had two main components: (1) Rehabilitation of the Lifeline Road Network and (2) Technical Assistance. The rehabilitation component included civil works for road rehabilitation, consultancy services for the supervision and technical auditing of rehabilitation works, and updating original MCA-financed designs and environmental documents to meet the requirements of the project. The technical assistance component featured technical assistance for strengthening ARD's capacity, including a study to review low-cost pavement options for Armenia, updated designs for a potential future project, and related training. This technical assistance component had a budget of less than $2 million.

Within eight months after the World Bank road rehabilitation project became effective, work crews had rehabilitated about 150 km of the LRN under 42 civil works contracts. Rehabilitation included paving all roads, implementing drainage systems, and introducing traffic safety

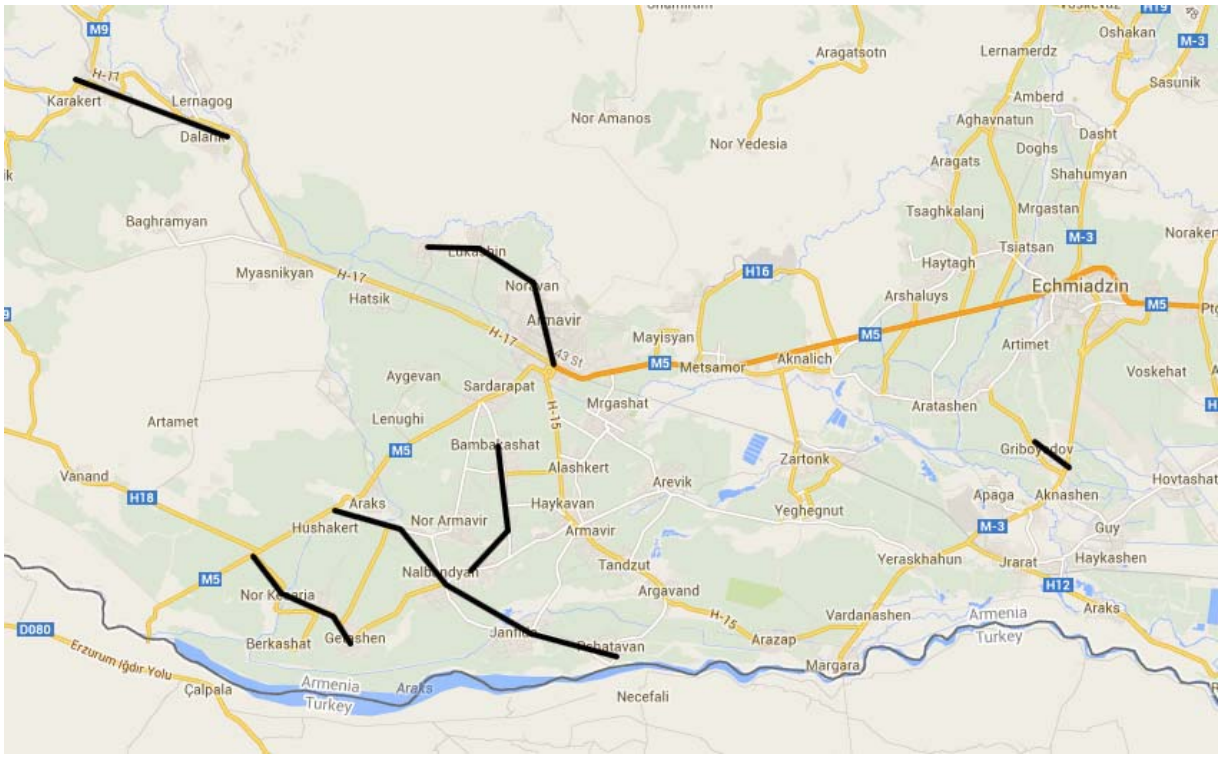
improvements. Contractors outperformed the project target by using cost savings to rehabilitate an additional 18 km of LRN.

On August 27, 2009, the World Bank approved additional financing of $36.6 million in loans to rehabilitate an additional 145.9 km of LRN and create a total of 10,000 person-months of new employment. The World Bank approved another $40 million in low-interest loans in July 2010, bringing the Bank's total lending for road rehabilitation to over $100 million. As of December 2013, 446 km of roads had been improved with World Bank funds (World Bank 2014). This met the project's final target of 430 kilometers. (See implementation timeline in Table II.1, and relationship with data collection in Figure). The Armenian government complemented this investment with a $16 million counterpart to rehabilitate an additional 50 km of the LRN in 2009 and 2010.

Table II.2 provides a summary of all road rehabilitation work financed by MCA, the World Bank, and the Republic of Armenia between 2007 and 2011. As illustrated, the World Bank made the largest investment in the form of long-term, low-interest loans. However, it should be noted that the Armenian government made large investments in rural roads as well, first with a $16 million investment in rehabilitating 50 km of the LRN, and again with over $25 million in counterpart investments and routine maintenance to interstate and lifeline roads per year.

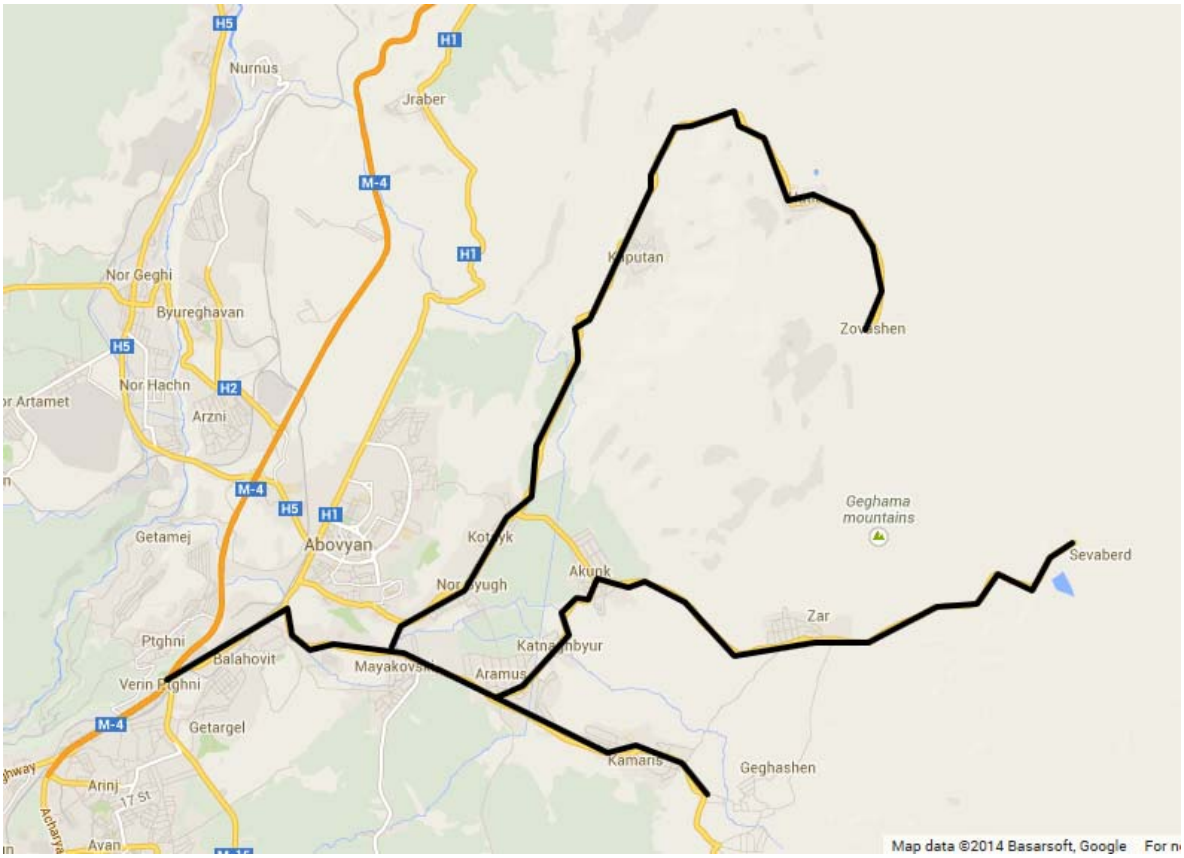
Figures II.1 and II.2 provide visual illustrations of the road links rehabilitated in Armavir and Kotayk, respectively, as examples of the types of roads rehabilitated under the project. As illustrated in the figures, rehabilitated roads generally connect relatively small villages to national highways (Highway M5 in the case of Armavir and Highway M4 in the case of Kotayk).

Figure II.1. Map of roads rehabilitated in Armavir, 2009-2010



Note: Rehabilitated roads are shown in black. Original map courtesy of Google Maps.

Figure II.2. Map of roads rehabilitated in Kotayk, 2009-2010



Note: Rehabilitated roads are shown in black. Original map courtesy of Google Maps.

Table II.2. Comparison of RRRP targets and outputs

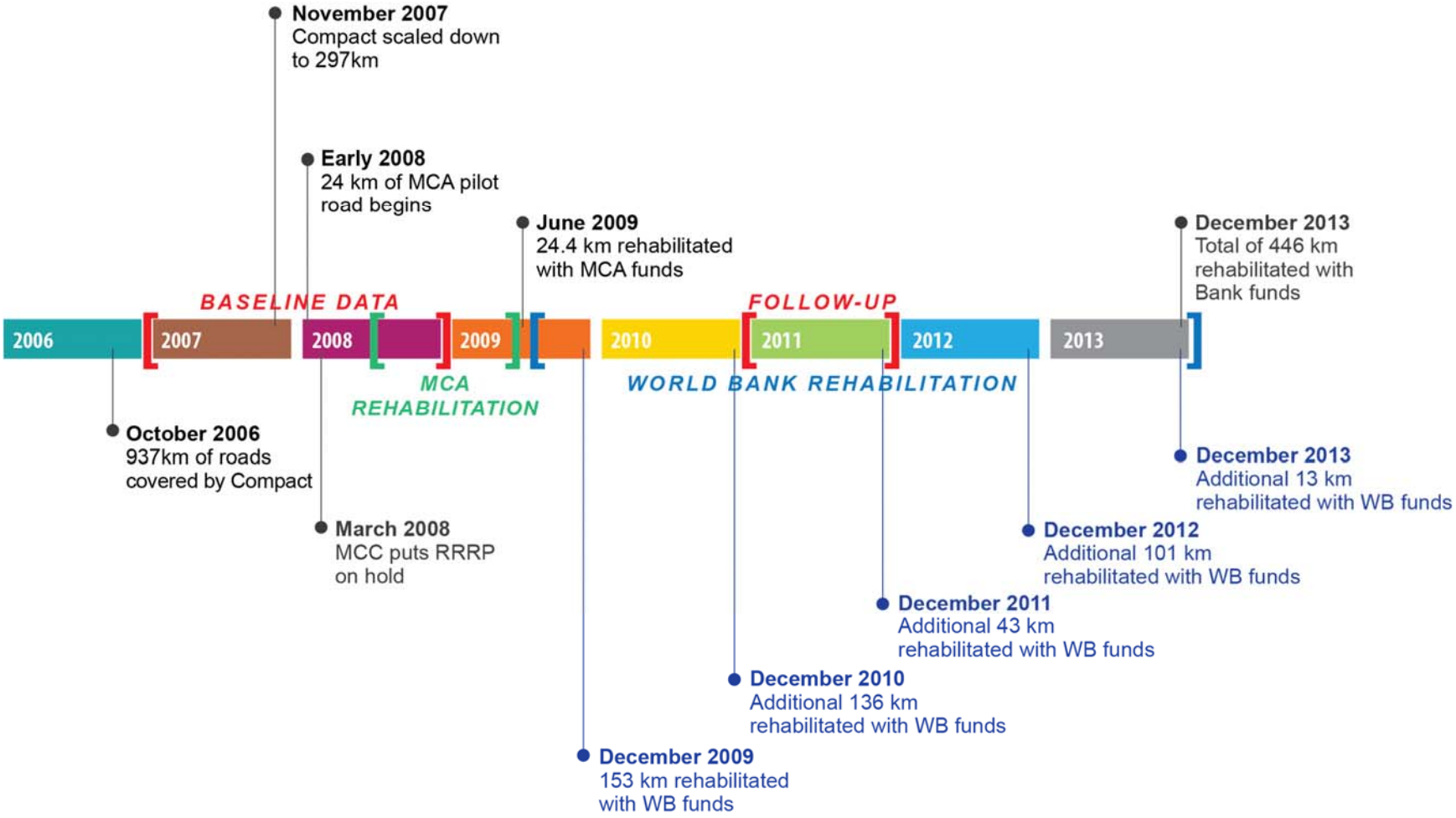
	Target	Actual
MCA		
Roads rehabilitated (in km)	943	24.4
Communities served	260	12
Number of beneficiaries	360,000	6,356
Investment (in millions)	$67	$8.4
World Bank: Round 1 (Apr 2009–Dec 2010)		
Roads rehabilitated (in km)	100	118
Investment (in millions)	$25	$25
RA (2009)		
Roads rehabilitated (in km)	50	50
Investment (in millions)	$16	$16
World Bank: Round 2 (Nov 2009–Dec 2011)		
Roads rehabilitated (in km)	140	146
Investment (in millions)	$36.6	$36.6
World Bank: Round 3 (Nov 2010–Dec 2013)		
Roads rehabilitated (in km)	190	182*
Investment (in millions)	$40	$40
World Bank: All rounds		
Roads rehabilitated (in km)	430	446
Investment (in millions)	$101.6	$101.6

Source: MCA-Armenia 2011 for MCA figures; World Bank 2013 for World Bank figures.

* Current as of 2014.

RA = Republic of Armenia

Figure II.3. RRRP implementation and data collection timeline



D. Implementation findings

MCA-funded rehabilitation. In addition to the impact estimates based on the ILCS, MCA-Armenia commissioned an independent qualitative process analysis (QPA) that examined the RRRP's design and implementation. The QPA is based on in-depth interviews with a small set of stakeholders, including program managers, construction firms, and community residents. The QPA focuses on learning how and why the RRRP was designed, the fidelity of the program as it was implemented, and stakeholders' perceptions of the strengths and challenges of the RRRP. The qualitative findings from the QPA provide valuable complementary information to help explain the quantitative findings from the impact evaluation. However, it is important to note that QPA findings were based exclusively on the construction for the 24.4 km of pilot roads funded by MCA-Armenia. According to MCA sources, pilot construction generally involved the same actors and activities as the large-scale rehabilitation efforts that would follow in 2009 and 2010. However, we caution against generalizing QPA findings on the pilot to the full RRRP given the small number of pilot roads relative to the full set of road links that were ultimately rehabilitated.

Overall, the QPA concluded that most of the outcome targets set for the RRRP in the MCA-Armenia Monitoring and Evaluation (M&E) Plan had been fully met on the two road sections that were rehabilitated with MCA funds. Particularly, as a result of project implementation, road roughness decreased, transportation costs decreased, and vehicular activity increased. The report also concluded that most of the activities envisioned under the RRRP had been effectively implemented; there were no major delays in project implementation, and the construction outputs were largely achieved. Notably, the QPA found that pilot rehabilitation efforts benefited from good planning and management.

World Bank-funded rehabilitation. Regarding the implementation of the World Bank–funded road rehabilitation project, program administrators judged the project to be a success. A document that summarized all World Bank-funded rehabilitation in 2009 stated that several key factors led to the successful implementation of the project. These included the Armenian government's full support for the project, a competent ARD implementation team, preexisting investments in selecting road links and completing construction designs, technical supervision provided by international firms, and local capacity and willingness to try new designs and technologies (World Bank 2010).

In particular, the summary report on World Bank–funded rehabilitation noted the capable ARD implementation team and the international consultant and World Bank technical auditor as key ingredients of success. Communication and interaction among partners characterized implementation. Throughout project implementation, the designer, supervision consultant, technical auditor, and contractors all collaborated under the leadership of ARD. Notably, supervisors and technical auditors were not just concerned about ensuring contractors' compliance. Rather, they offered regular advice on how to improve quality and provided hands-on guidance, thus transferring technology and knowledge to contractors, designers, and ARD. In addition, ARD project managers regularly visited contractors to oversee their work. Stakeholders noted that ARD staff was willing to adopt new designs and technologies. In addition, well-staffed Bank supervision missions and an independent technical auditor provided advice to the client on all major aspects of the project (World Bank 2010).

E. Program results

MCA-funded rehabilitation. Road roughness dramatically decreased following rehabilitation from the MCA-funded pilot RRRP. According to the MCA-Armenia M&E Plan, the baseline (2007) for the International Roughness Index (IRI)[20]—which uses a mechanical device to measure the relative deviation of a vehicle's rear suspension from its bottom—was 14.2 m/km. Stakeholders identified an IRI of 4 m/km among pilot roads as a result of the intervention. The RRRP actually surpassed this target; an IRI measurement conducted by the ARD after the rehabilitation of the pilot roads indicated an IRI of 3.5 m/km (MCC 2012).

According to the MCA-Armenia M&E Plan, the average daily traffic on pilot roads was 637 vehicles in 2007 before the project (Table II.3). Following road rehabilitation, this number increased to 735, according to average daily traffic counts on pilot roads conducted by ARD (ARD 2008). We note, however, that we cannot necessarily attribute this increase in traffic to the road rehabilitation project without having an estimate of how much traffic would have changed in the absence of rehabilitation. Traffic could have increased (or decreased) due to other factors outside the RRRP.

Table II.3. RRRP key performance indicators, MCA-funded RRRP

	Baseline	Target	Actual (Sept. 2011)
International Roughness Index (lower number = smoother road)	14.2	4	3.5
Average annual daily traffic on pilot roads (in vehicles)	637	706	735

Source: Table of Key Performance Indicators, Armenia. MCC 2012 and ARD 2008.

As mentioned above, MCC estimated the ERR of the original project at 26 percent over 30 years. In 2011, MCC estimated revised ERRs that reflected only the costs and benefits of the RRRP pilot roads, which totaled approximately 24.4 km. According to updated ERR calculations, the RRRP portions funded by MCC had an ERR of 18 percent over 20 years; this is about 8 percentage points less than the original ERR but still well above the original threshold of 12.5 percent (MCC 2011).[21]

As documented in the QPA, project beneficiaries of the RRRP were mostly satisfied with the project's results. The respondents noted reduced transportation costs, reduced time spent traveling, increased driving speed, and a reduced number of emergency situations/car accidents caused by bad conditions of the road. According to interviewed stakeholders, rehabilitation resulted in better access to public and private transportation, and generated agricultural benefits related to reduced transportation costs. In particular, farmers in some beneficiary communities

[20] According to the M&E plan, the outcome of reduction of transportation costs is measured by the International Roughness Index developed by the World Bank. Data are collected by the RoadMaster Bump Integrator method. Measurements are made by the RoadMaster device, which is attached to the rear wheel of a car.

[21] It does not appear that final ERRs used updated IRI and traffic count data. Rather, the model used original assumptions based on population counts and car ownership in communities served by the pilot, and updated costs and benefits to reflect the smaller scope of the MCC-funded portion of the RRRP.

reported elevated income as a result of increased competition among intermediaries, who were more likely to travel to rural communities as a result of completed rehabilitation.

The QPA also noted that temporary employment rose in beneficiary communities. Some residents of beneficiary communities were hired for short-term construction works as part of the project. Residents involved in construction were highly satisfied with their participation, mostly because they had a comparatively stable and well-paid job. In addition, in half of the beneficiary communities visited by researchers, women reported being indirectly involved in road construction, providing food and cleaning services to road crews (Socioscope 2010).

World Bank–funded rehabilitation. The World Bank staff primarily tracked the contribution of the roads rehabilitation projects to temporary employment. As illustrated in Table II.4, the project generated nearly 40,000 person-months of employment from 2008 to 2013. In addition, the project dramatically surpassed the target of a 20 percent reduction in travel time on rehabilitated road links, with an average decrease of 59 percentage points in travel time across all links and an average reduction in transport costs of 26 percent (World Bank 2014).[22]

Table II.4. Employment in World Bank–funded rehabilitation, 2009–2013

	Target	Actual (June 2013)
Person-months of employment created: all phases (2009–2013)	36,650	39,855
Reduction in travel time	20%	58.5%
Reduction in transport costs	20%	25.8%

Source: World Bank 2014.

Note: All figures current as of December 2013.

F. Evaluation sample

Table II.5 shows the demographic and economic characteristics of households in the analysis sample. Because the sample is a repeated cross-section of different households interviewed each year, data for 2007 and 2008 are presented in separate tables. Treatment and comparison groups had similar demographic characteristics at both baseline years, with the exception of head-of-household educational attainment and age. Head-of-household educational achievement was evenly matched in 2008. But in 2007, comparison households had, on average, a higher education attainment relative to treatment households. Forty-four percent of heads of household in the comparison group completed a full secondary education, relative to 34 percent in the treatment group. In 2008, heads of household in the treatment group were older, on average, than in the comparison group (61 years and 59 years, respectively). About 30 percent of households in both 2007 and 2008 reported a female head, which was not statistically significantly different for the treatment and comparison groups.

On average, the baseline income and consumption of treatment and comparison households are somewhat similar but not perfectly matched. In both 2007 and 2008, comparison households

[22] Reductions in transport costs were estimated using the HDM-4 model based on measured roughness values.

had significantly higher average income from food sales ($480 vs. $195, and $472 vs. $298, respectively). Comparison households also had higher average income from crop sales in both years, although this difference was only significant in 2008 ($691 vs. $357). On average, comparison households' total household income was greater than treatment households' total income in 2007 ($2,869 vs. $2,439), although this difference is not statistically significant. In 2008, this difference was less than $160 and was not statistically significant. The observed treatment-comparison differences in household income translate into differences in consumption of similar magnitudes, with comparison group households reporting higher consumption than treatment households. See distributions of treatment and comparison group consumption at baseline in figures in Appendix A.

In addition, treatment households had lower poverty rates than comparison households in 2007 (13 percent in poverty versus 20 percent in comparison), despite also having lower household income on average because of differences in their distributions, with the comparison households having somewhat higher percentages of both higher- and lower-income households. This stands in contrast to poverty rates in 2008, which were similar for the treatment and comparison groups. Because these outcomes of income, consumption, and poverty are crucial in the program logic, it is especially important that our evaluation design adequately controls for these baseline differences.

Table II.5. Demographic and economic characteristics at baseline, 2007 (percentages unless otherwise indicated)

	2007				2008			
Characteristic/Measure	T	C	Diff	p-Value	T	C	Diff	p-Value
Demographic Characteristics								
Head of Household's Age (years)	60.5	59.3	1.2	0.31	61.4	58.8	2.7**	0.05
Female-Headed Household (%)	32	29	3	0.39	29	28	0	0.89
Head of Household's Education								
Less than secondary	38	32	6	0.16	80	79	1	0.71
Full secondary	34	44	-10**	0.02	14	17	-3	0.23
Secondary vocational	7	5	3	0.29	0	0	0	0.52
More than secondary	21	20	1	0.79	6	4	2	0.32
Total People in Household	4.4	4.5	-0.1	0.52	4.4	4.3	0.1	0.59
Number of Children in Household	1.2	1.2	0.0	0.73	1.2	1.2	0.0	0.77
Total Land Owned or Rented (Hectares)	1.9	2.0	-0.1	0.81	1.5	1.4	0.1	0.75
Number of Animals Owned:								
Cows	1.1	1.4	-0.3	0.26	1.0	1.3	-0.3	0.16
Sheep	1.6	2.5	-0.9	0.35	1.2	1.7	-0.5	0.34
Pigs	0.2	0.3	-0.1	0.19	0.1	0.2	-0.1	0.11
Household Income, Consumption, and Poverty (US$)								
Income from Crop Sales	285	497	-212	0.19	357	691	-334*	0.07
Income from Food Sales	195	480	-285***	0.00	298	472	-174*	0.08
Income from Animal Sales	117	174	-57	0.28	157	206	-49	0.38
Total Household Income	2,439	2,869	-430	0.13	3,156	3,312	-157	0.57
Annual Household Consumption	3,423	3,712	-289	0.42	4,258	4,278	-20	0.96
Poor (%)	13	20	-8*	0.08	16	17	0	0.90
Households	848	832			880	680		
Road Links	26	26			23	27		

Source: 2007 and 2011 ILCS.

Note: All dollars are in constant 2011 dollars, converted from Armenian drams using the OANDA currency converter and into constant 2011 dollars using the U.S. GDP deflator.

*/**/*** statistically significant at 10, 5, and 1 percent, respectively

T = Treatment; C = Comparison; Diff = Difference between Treatment and Comparison Means.

III. IMPACTS OF THE RRRP

In this chapter, we present estimated RRRP impacts on immediate, short-term, medium-term, and long-term impacts. As was explained in Chapter I, we reiterate that many impacts are imprecisely estimated due to the relatively small number of roads that are included in the analysis. This is especially important when interpreting the estimates on highly variable outcome measures such as income. As a result, we interpret some findings as suggestive rather than conclusive.

A. Expected pathway of RRRP impacts

Before discussing program impacts, it is useful to review how the RRRP was designed to generate the Compact's ultimate goal of reduced poverty. Improved rural roads would reduce travel time and vehicle operation costs (immediate outcomes), which would enhance residents' access to markets and social infrastructure (short-term outcome). As a result of decreased transportation costs, residents could access inputs at a cheaper price and make other long-term investments, which would in turn boost production (medium-term outcomes). Also due to improved roads, an increased number of retailers and collectors of agricultural products could access the communities, thus creating conditions for farmers to sell a larger share of their agricultural production at a better price (medium-term outcome). Better road infrastructure would also result in increased investment in beneficiary communities and employment opportunities for residents (medium-term outcomes), which would improve household income and consumption, and decrease poverty rates (long-term outcomes).

In theory, immediate outcomes of reduced travel time and costs could materialize directly following rehabilitation, short-term outcomes of enhanced access could commence in the months following rehabilitation, medium-term outcomes of increased investment and production could occur within a year of rehabilitation but more likely two or more years, and long-term impacts of increased household income and reduced poverty would be expected to materialize within several years of the intervention.

This program logic forms the basis for our impact analysis in that we first analyze impacts of the RRRP on immediate outcomes and short-term outcomes, and proceed to analyze impacts of the intervention on medium- and long-term outcomes. Reviewing these outcomes in temporal order helps us examine and validate key linkages in the program logic in the order in which they were envisioned, and identify the point at which the program logic breaks down—or fails to provide rigorous evidence of program impacts.

This exercise also helps identify areas in which time constraints may be responsible for a lack of impacts on medium- and long-term outcomes. Particularly relevant for this impact analysis, treatment households experienced road improvements between one and two years prior to the follow-up survey. As such, the follow-up period for this evaluation is relatively short—as little as one year for some road links—and it is earlier than we would expect long-term impacts to materialize. We nonetheless examine impacts on long-term impacts to see if there are any early signs of impacts, but an absence of impacts on income, consumption, and poverty rates should not be taken as an indication that the program failed. Given the follow-up period, however, we would expect to see improvements in short-term outcomes, such as perceived road

quality, and perhaps in some medium-term outcomes, such as agricultural investments. If short-term impacts are evident in this analysis of 2011 data, there is some possibility that medium- and longer-term impacts—such as increased household income and consumption—may occur in the future.

In this analysis, we define an estimated impact as the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups, as described in Chapter II. Our findings are robust to alternative specifications to account for baseline differences between the two groups, as described in Appendix A.

B. Impact of rehabilitation on quality, access, and utilization

Regarding immediate and short-term outcomes of road quality and access, we found statistically significant impacts on households' perceptions of road quality and their use of road links (the top panel of Table III.1). Treatment households were 39 percentage points more likely than households in the comparison group to rate regional roads—the type of roads that were rehabilitated—as good or excellent (p-value: < 0.01). There is no impact on ratings of local roads, which were not rehabilitated. We also observe positive, statistically significant impacts on the likelihood of households reporting transportation services as good or excellent (19 percentage points; p-value: 0.05).

The impact of road rehabilitation on reported problems with market access was also large and statistically significant. Treatment households were 20 percentage points more likely than households in the comparison group to report no problems with market access during a typical month, and this estimate is statistically significant (p-value: < 0.01). We also observed a positive and statistically significant impact on the likelihood of using roads for a purpose other than getting to work or buying and selling agricultural products during a typical month (17 percentage points; p-value: 0.01). These "other" purposes typically included shopping or visiting relatives.

Table III.1. Impact of RRRP on perceived quality and road utilization (averages unless otherwise indicated)

Measure	Treatment Mean	Comparison Mean	Program Impact	p-Value of Impact
Percentage Reporting:				
Local Roads Are Good or Excellent	5	2	3	0.17
Regional Roads Are Good or Excellent	52	12	39***	<0.01
Transportation Services Are Good or Excellent	41	22	19*	0.05
No Market Access Problems	91	71	20***	<0.01
Percentage That Use Roads for:				
Buying Agricultural Inputs	15	6	9*	0.10
Selling Agricultural Products	18	26	-8	0.20
Getting to Work	5	2	2	0.36
Other Purposes	88	71	17**	0.01

Measure	Treatment Mean	Comparison Mean	Program Impact	p-Value of Impact
Days a Month Roads are Used for:				
Buying Agricultural Inputs	0.3	0.2	0.0	0.72
Selling Agricultural Products	1.2	1.8	-0.6*	0.09
Getting to Work	2.2	3.0	-0.8	0.24
Other Purposes	3.6	3.0	0.6	0.19
Households	2,560	2,288		
Road Links	27	28		

Source: 2007, 2008, and 2011 ILCS.

Notes: Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place. Questions reflect respondents' estimates for a typical month.

*/**/*** statistically significant at 10, 5, and 1 percent, respectively

In addition, treatment households were 9 percentage points more likely than comparison households to report using roads to buy agricultural inputs (p-value: 0.10; Table III.1). Treatment households also reported spending less time on roads to sell agricultural products (treatment-control difference of 0.6 days per month). This last impact may reflect the increased use of rehabilitated roads by local agriculture wholesalers, who may have traveled to treatment group households' farm plots to buy their production, thus saving these households a trip to the local market. This scenario is consistent with the RRRP logic model and QPA findings in Chapter 3 that agricultural intermediaries were more likely to travel to rural communities as a result of completed rehabilitation. Another feasible possibility is that rehabilitated roads cut down on travel time, to the extent that treatment households spent significantly less time in transit.

Examining short-term impacts by gender of the head of household, we see that male-headed households were more likely than female-headed households to report improved market access (Table III.2). However, female-headed households were more likely to report using roads for other purposes. Related to this finding, female-headed households appeared to use roads approximately one more day per month to visit relatives, go shopping, and for other non-income generating activities as a result of the RRRP. This positive impact was not present among male-headed households.

Table III.2. Impact of road rehabilitation on perceived quality and road utilization, by gender of head of household (averages unless otherwise indicated)

Measure	Overall Impact: All Households	Program Impact: Female-Headed Households	Program Impact: Male-Headed Households
Percentage Reporting:			
Local Roads Are Good or Excellent	3	3	3
Regional Roads Are Good or Excellent	39***	45***	37***
Transportation Services Are Good or Excellent	19*	17	21**
No Market Access Problems	20***	12**	23***
Percentage That Use Roads for:			
Buying Agricultural Inputs	9*	11	8
Selling Agricultural Products	-8	-3	-10
Getting to Work	2	3	2
Other Purposes	17**	23***	14*
Days a Month Roads are Used for:			
Buying Agricultural Inputs	0.0	0.0	0.1
Selling Agricultural Products	-0.6*	-0.6	-0.6
Getting to Work	-0.8	-0.5	-0.9
Other Purposes	0.6	1.1*	0.4
Households	4,731	1,346	3,385
Road Links	55	55	55

Source: 2007, 2008, and 2011 ILCS.

Notes: Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place.

*/**/*** statistically significant at 10, 5, and 1 percent, respectively

Examining the short-term outcomes of access to social infrastructure and utilization, we found few statistically significant impacts (Table III.3). There were no significant impacts on the likelihood of using a car or bus, nor were there statistically significant impacts on the likelihood of using a car or bus to get to the hospital and pharmacy. It should be noted that treatment households were 12 percentage points more likely than comparison households to report using a car or bus to travel to the hospital or pharmacy. However, the p-values for these differences of 0.23 and 0.28, respectively, preclude confirmation that the RRRP allowed a larger portion of households to use roads for these purposes. In addition, residents' distance in minutes to the hospital, pharmacy, and other social infrastructure was similar among treatment and comparison group households.

Table III.3. Impact of road rehabilitation on utilization and access to social infrastructure (averages unless otherwise indicated)

Measure	Treatment Mean	Comparison Mean	Program Impact	p-Value of Impact
Percentage Reporting:				
Using a Car	60	57	2	0.81
Using a Bus	73	74	-1	0.91
Percentage Reporting Using a Car or Bus to Get to the:				
Hospital	71	58	12	0.23
Pharmacy	63	50	12	0.28
Community Center	4	5	0	0.91
Kindergarten	37	42	-5	0.67
Secondary School	4	7	-3	0.46
Distance in minutes to:				
Hospital	19.5	18.8	0.7	0.88
Pharmacy	16.2	19.3	-3.2	0.55
Community Center	11.1	12.8	-1.7	0.30
Kindergarten	17.6	15.7	1.9	0.57
Secondary School	11.8	12.7	-0.8	0.63
Households	2,560	2,288		
Road Links	27	28		

Source: 2007, 2008, and 2011 ILCS.

Notes: Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place.

Regarding medium-term outcomes, we found little evidence of impacts on agricultural investments (Table III.4). Treatment households spent more on hired labor than did comparison households, nearly double the comparison mean of $24. This estimated impact is on the margin of statistical significance, but most households spent little on any hired labor, and this estimated impact is not economically meaningful. There are no other notable impacts on agricultural spending.

Table III.4. Impact of road rehabilitation on investments (averages unless otherwise indicated)

	Treatment Mean	Comparison Mean	Program Impact	p-Value of Impact
Total Agricultural Investments (in US$)	517	459	58	0.29
Irrigation	44	43	1	0.88
Seeds and seedlings	35	33	2	0.88
Fertilizers	37	41	-5	0.58
Herbicides	34	37	-4	0.66
Hired labor	45	24	21	0.12
Transportation	44	45	-2	0.90
Fodder	152	112	39	0.19
Taxes and duties	32	39	-7	0.17
Rented equipment	61	51	10	0.55
Other expenses	34	32	2	0.85
Households	2,560	2,288		
Road Links	27	28		

Source: 2007, 2008, and 2011 ILCS.

Note: All dollars are in 2011 dollars, converted from Armenian Drams using the OANDA currency. Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place.

Regarding additional medium-term outcomes, no statistically significant impacts were found for agricultural production or sales (Tables III.5 and III.6). In Table III.7, we present the estimates of food production and sales. Examining the production of individual foods, only the 5 percentage point impact on the likelihood of jam production (p-value: 0.03) and the 18.6 kg impact on the amount of preserved vegetables produced (p-value: 0.01) were statistically significant. Weighed against these positive impacts, treatment households had lower egg sales than comparison households (negative impact of $31 in egg sales; p-value of 0.10) and got a slightly lower price when selling eggs and grass. Overall, total incomes from food sales were almost identical for the treatment and comparison groups, and there were no statistically significant impacts on sales for any specific foods. Similarly, conditional on selling a given crop or food product, there were no meaningful differences in the prices received by treatment or comparison households.

Analyzing these medium-term outcomes by gender of the head of household (not shown in a separate table), we found an absence of positive impacts on the value of crop and food sales for both male- and female-headed households. The only noteworthy negative impact of the RRRP on crop and food sales—the loss of $31 in egg sales mentioned above—appears largely driven by male-headed households, which experienced a negative impact of $40 in annual egg sales (significant at 10 percent). However, this negative impact is not large in magnitude, particularly for sales over the course of the entire calendar year.

Table III.5. Impact of road rehabilitation on agricultural production (averages unless otherwise indicated)

Measure	Treatment Mean	Comparison Mean	Program Impact	p-Value of Impact
Percentage Cultivating				
Fruits	58	71	-14	0.12
Vegetables	71	71	0	0.99
Potatoes	44	41	3	0.70
Grapes	23	18	5	0.41
Grains	21	20	1	0.89
Grass	37	40	-3	0.67
Beans	19	17	2	0.80
Production (kg)				
Fruits	342.3	390.8	-48.5	0.45
Vegetables	1,137.2	1,211.1	-73.9	0.72
Potatoes	541.1	602.0	-60.8	0.77
Grapes	550.2	681.7	-131.6	0.48
Grains	518.0	594.9	-76.9	0.81
Grass	2,781.0	2,992.0	-211.0	0.89
Beans	16.9	3.9	13.0	0.11
Households	2,560	2,288		
Road Links	27	28		

Source: 2007, 2008, and 2011 ILCS.

Note: All dollars are in 2011 dollars, converted from Armenian Drams using the OANDA currency. Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place.

Table III.6. Impact of road rehabilitation on agricultural sales (averages unless otherwise indicated)

Measure	Treatment Mean	Comparison Mean	Program Impact	p-Value of Impact
Percentage Selling				
Fruits	24	24	1	0.89
Vegetables	31	29	2	0.74
Potatoes	23	22	1	0.93
Grapes	12	13	-1	0.82
Grains	13	8	5	0.37
Grass	9	5	4	0.29
Beans	6	4	2	0.64
Sales (kg)				
Fruits	197.2	245.2	-48.0	0.20
Vegetables	928.8	971.0	-42.2	0.83
Potatoes	339.5	311.1	28.4	0.85
Grapes	465.3	635.6	-170.4	0.30
Grains	454.0	405.5	48.5	0.84
Grass	263.4	244.8	18.6	0.92
Beans	13.1	1.5	11.6	0.12
Sale Price ($/kg)[a]				
Fruits	1.26	1.19	0.07	0.83
Vegetables	0.44	0.54	-0.11	0.32
Potatoes	0.49	0.47	0.02	0.58
Grapes	0.53	0.49	0.04	0.28
Grains	0.46	0.46	0.00	0.96
Grass	0.07	0.09	-0.02**	0.04
Beans	2.30	2.47	-0.17	0.34
Sales (US$)				
Fruits	142	175	-33	0.33
Vegetables	278	371	-92	0.22
Potatoes	140	142	-2	0.98
Grapes	206	254	-47	0.53
Grains	91	89	2	0.98
Grass	39	33	6	0.73
Beans	26	4	23	0.14
Total income from crop sales	1,000	1,100	-99	0.60
Households	2,560	2,288		
Road Links	27	28		

Source: 2007, 2008, and 2011 ILCS.

Note: All dollars are in 2011 dollars, converted from Armenian Drams using the OANDA currency. Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place.

*/**/*** statistically significant at 10, 5, and 1 percent, respectively

[a] Price received per kilogram or liter is defined only for households that sold the specified crop

Table III.7. Impact of road rehabilitation on food production and sales (averages unless otherwise indicated)

Measure	Treatment Mean	Comparison Mean	Program Impact	p-Value of Impact
Percentage Producing				
Milk	46	48	-3	0.61
Cheese	46	47	-1	0.81
Sweet syrup	92	89	3	0.26
Jam	93	88	5**	0.03
Preserved Vegetables	92	87	5	0.10
Bread	86	83	2	0.46
Eggs	61	56	5	0.47
Production (kg unless otherwise noted)				
Milk (in liters)	1,129.0	875.1	253.9	0.43
Cheese	63.8	72.7	-8.9	0.71
Sweet syrup	81.5	87.1	-5.6	0.54
Jam	25.3	20.6	4.6	0.10
Preserved Vegetables	67.8	49.2	18.6**	0.01
Bread	487.4	460.6	26.8	0.39
Eggs	797.9	890.7	-92.8	0.55
Percentage Selling				
Milk	43	34	8	0.20
Cheese	30	26	5	0.48
Eggs	25	21	3	0.63
Yogurt	14	19	-5	0.13
Butter	9	6	2	0.63
Sold (kg)				
Milk (in liters)	1,010	765	244	0.44
Cheese	31	43	-12	0.56
Eggs	212	363	-151	0.17
Yogurt	51	67	-16	0.28
Butter	3	1	2	0.15
Sale Price (US$/kg or liter) [a]				
Milk	0.44	0.41	0.03	0.25
Cheese	3.58	3.75	-0.17	0.45
Eggs	0.14	0.15	-0.01**	0.03
Yogurt	0.57	0.60	-0.03	0.27
Butter	5.33	5.49	-0.17	0.54
Sold (US$)				
Milk	355	266	89	0.39
Cheese	95	151	-56	0.40
Eggs	31	62	-31*	0.10
Yogurt	32	41	-9	0.35
Butter	17	9	8	0.20
Total income from food sales	561	549	12	0.92
Households	2,560	2,288		
Road Links	27	28		

Source: 2007, 2008, and 2011 ILCS.

Note: All dollars are in 2011 dollars, converted from Armenian Drams using the OANDA currency. Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place.

*/**/*** statistically significant at 10, 5, and 1 percent, respectively

[a] Price received per kilogram or liter is defined only for households that sold the specified crop

Regarding animal purchases and ownership, we observed positive, significant impacts on the number of sheep owned (1.3, p-value: 0.08). We also estimate positive impacts on the number of cows bought (less than 0.1; p-value: 0.10) and the number of pigs bought (0.1; p-value: 0.05), but both treatment and comparison households bought very few of both types of livestock. In addition, we found no conclusive impacts on total income from animal sales between treatment and comparison groups (Table III.8). However, treatment households appeared to spend slightly more, on average, on animal purchases than comparison households (impact of $31, p-value of 0.06).

Integrating findings regarding all types of agricultural investments, production, and income, there is little evidence that suggests the RRRP has had an overall impact on households' agricultural production or income. There are a few positive impacts—such as production of jam and vegetable preserves—but there are also a few negative impacts, such as egg sales. These statistically significant impacts are relatively small, and considering the large number of outcomes that were examined, this handful of significant impacts (both positive and negative) is quite possibly due to chance. As noted previously, the follow-up period was probably too short for us to expect impacts on production or income to have materialized, and a lack of statistical power further complicates our ability to discern the RRRP's impact. On the other hand, there is also no evidence of impacts on most agricultural investments within a one- to two-year time frame. This suggests that households have not yet changed their investment and production patterns in response to road improvements, or at least had not done so at the time of the 2011 survey. Similarly, there is no evidence of changes in prices, which could have been expected to have begun by this point.

There is some evidence, albeit inconclusive, that households may have increased investments and output of some animal products. Treatment households are more likely to buy cows and sheep, appear to spend more on animal fodder and hired labor, and appear more likely to sell milk than comparison households. It is possible that improved roads played a role in expanding treatment households' access to animals, hired labor, and animal fodder, as well as their access to milk consumers. However, given the lack of statistical significance of these treatment-comparison differences, we cannot conclude that road rehabilitation investments generated increased milk production among treatment households.

Table III.8. Impact of road rehabilitation on animal ownership and sales (averages unless otherwise indicated)

Measure	Treatment Mean	Comparison Mean	Program Impact	p-Value of Impact
Number of cows:				
Owned	1.2	1.1	0.1	0.75
Bought	0.0	0.0	<0.0*	0.10
Sold	0.0	0.0	0.0	0.46
Number of pigs:				
Owned	0.5	0.3	0.2	0.11
Bought	0.1	0.0	0.1*	0.05
Sold	0.5	0.4	0.1	0.24
Number of sheep:				
Owned	2.8	1.5	1.3*	0.08
Bought	0.0	0.0	0.0	0.35
Sold	0.7	0.6	0.2	0.49
Total expenditure on animal purchases	37	5	31*	0.06
Total income from animal sales	327	375	-48	0.59
Total income from meat sales	90	112	-22	0.39
Households	2,560	2,288		
Road Links	27	28		

Source: 2007, 2008, and 2011 ILCS.

Note: All dollars are in 2011 dollars, converted from Armenian Drams using the OANDA currency. Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place.

*/**/*** statistically significant at 10, 5, and 1 percent, respectively

C. Impact of RRRP on household income

Total household 2011 income was not significantly different between treatment and comparison households, nor were there impacts on specific sources of income (Table III.9). Though the impact estimates are all statistically insignificant, the magnitudes of some estimates are large. In particular, the estimated impact of −$431 on total household income is 12 percent of the comparison households' average income of $3,653, and the estimated impact of −$348 on remittances is about 60 percent of the comparison households' average remittances. It should be noted that income is highly variable in the data, especially income from remittances, and as our minimum detectable impacts indicate, only large impacts on income can be reliably detected. Because of the short follow-up period for this evaluation, impacts on income are not a priority of the present report, but the reality of imprecise impact estimates for household income should be borne in mind if a future evaluation examines longer-term impacts.

Table III.9. Impact of road rehabilitation on annual household income (averages in US$ unless otherwise indicated)

	Treatment Mean	Comparison Mean	Program Impact	p-Value of Impact
Total household income	3,222	3,653	-431	0.29
Employment income	1,253	1,453	-200	0.37
Agricultural income	810	756	54	0.81
Remittances	199	547	-348	0.13
Pensions	697	706	-9	0.87
Other sources	263	191	72	0.21
Households	2,560	2,288		
Road Links	27	28		

Source: 2007, 2008, and 2011 ILCS.

Note: All dollars are in 2011 dollars, converted from Armenian Drams using the OANDA currency. Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place.

As described in Chapter I, one of the key goals of the World Bank–funded portion of the RRRP was to increase short-term employment by employing local workers in the rehabilitation projects. To test whether the project did in fact generate short-term employment in beneficiary communities, we also examined impacts of the RRRP on employment income in the years in which projects were implemented. Roads were rehabilitated in three distinct years—2009, 2010, and 2011—but there were not enough roads rehabilitated in any single year to have meaningfully precise estimates for any year. In contrast to the previous impact estimates that focus on impacts measured only in 2011, we instead pool data from the three years to measure impacts on short-term employment income. To do this, we measured employment income from 2009 if the household was served by a road link rehabilitated in 2009, and likewise for road links rehabilitated in 2010 and 2011. This specification also includes roads that were rehabilitated in 2011, whereas all other models only included road links rehabilitated in 2009 or 2010. The regression model is otherwise identical to the model used for other estimates.

Using this approach, we find no evidence of impacts on short-term employment (Table III.10). Employment income was the same, on average, for treatment households in the years in which their roads were rehabilitated as for comparison households in those same years. There are several possible explanations for this. One possibility is that household members employed in rehabilitation projects substituted this construction work for other employment that they would have otherwise had, including working abroad in Russia, with no change in their net income. Another possible explanation is that employment generated by the RRRP, estimated as 10,000 person-months during this time frame, was small relative to the population in these communities, thus diluting the per-household impact that we measure using ILCS data. A third possible explanation is that many of the households whose members were employed by the RRRP do not reside in the communities that were served directly by the RRRP, and thus are not covered by the ILCS analysis sample. All of these explanations may contribute to some degree, but we unfortunately cannot distinguish between them or assess their relative contributions to the null impact estimate. We note, however, that whereas nearly all of the other impact estimates were robust to alternative empirical specifications, the estimated impacts on short-term employment

were not. In one of the five alternative specifications, the estimated impact was positive and marginally statistically significant. We cannot say conclusively that there was no impact on short-term employment.

Table III.10. Impact of road rehabilitation on annual household income (averages in US$ unless otherwise indicated)

	Treatment Mean (Roads Rehabilitated in Given Year)	Comparison Mean	Program Impact	p-Value of Impact
Employment Income in Rehabilitation Year, Pooled	1,228	1,226	2	0.99
Households	2,968	3,680		
Road Links	33	28		

Source: 2007, 2008, 2009, 2010, and 2011 ILCS.

Note: All dollars are in 2011 dollars, converted from Armenian Drams using the OANDA currency. Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place.

Estimated impacts on household income do not differ much for female- and male-headed households (Table III.11), either for total income or specific types of income. The magnitudes of the impact estimates are similar for both types of households, and none are significant at conventional levels except for impacts on remittances. For remittances, female-headed households have an estimated impact of −$604, which is significant at the 10 percent level. There is no self-evident hypothesis regarding how a road rehabilitation project could affect remittances to female-headed households. Most likely, this negative impact of the RRRP on remittances reflects the increased probability of Type 1 errors, or false positives in statistical hypothesis testing, associated with examining impacts among a large number of outcomes.

Table III.11. Impact of road rehabilitation on annual household income, by gender of head of household (averages in US$ unless otherwise indicated)

	Overall Impact: All Households	Program Impact: Female-Headed Households	Program Impact: Male-Headed Households
Income from crop sales	-99	-245	-22
Income from food sales	12	-33	46
Income from animal sales	-48	64	-89
Total household income	-431	-549	-334
Employment income	-200	-151	-184
Remittances	-348	-604*	-248
Agricultural income	54	122	39
Pensions	-9	-10	3
Other sources	72	94	56
Households	4,731	1,346	3,385
Road Links	55	55	55

Source: 2007 and 2011 ILCS.

Note: All dollars are in 2011 dollars, converted from Armenian Drams using the OANDA currency. Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place.

*/**/*** statistically significant at 10, 5, and 1 percent, respectively

D. Impact of rehabilitation on consumption and poverty

In the long term, the RRRP was intended to increase income, which would in turn increase household consumption and reduce poverty. Consumption-based measures of household well-being are also a useful complement to income because consumption is less variable from year to year than income (Deaton 1997). As with the impacts on income, the follow-up period for this evaluation was too short for us to expect the full impacts on consumption and poverty to have materialized. However, we examine impacts on consumption and poverty to determine if there is evidence of early impacts on these key outcomes.

Poverty status is calculated by the National Statistical Service of Armenia based on household consumption, which itself is based on detailed household consumption diaries. Respondents record daily purchases and consumption of most food and non-food items over the course of a month. Items that are consumed regularly but in small amounts, such as spices or salt, are recorded on a monthly basis. Purchases of durable goods, such as vehicles and appliances, are reported for the previous 12 months.[23] Using the diary responses, the reported value of all consumption is annualized and summed for the household, including food, health care, other nondurable goods, and durable goods. This sum is then adjusted based on the number of adults and children in the household to determine consumption per person. Then, the estimate of total consumption per person is compared to the national poverty line, which is defined as the monetary value of the minimum consumer basket. This represents the amount of goods and services that meet the needs of the minimum level of Armenian living standards. The minimum

[23] The instruments can be accessed at http://www.armstat.am/en/?nid=378.

consumer basket consists of two components: (1) a minimum food basket and (2) a set of basic non-food goods and services.

In 2009, the ILCS instituted a series of methodological changes to improve the accuracy of its poverty measures. For example, the new questionnaire takes into account a greater variety of food items and the exact number of days each household member in its survey was present in the household. As part of this new methodology, the former metric for poverty, based on the "complete poverty line," was replaced with two poverty lines—the "upper poverty line," which assumes food consumption is 56.5 percent of total consumption, and the "lower poverty line," which estimates food consumption as 70 percent of all household consumption. The changed definition of poverty is potentially problematic for our difference-in-differences estimation strategy, which assumes outcomes are measured in the same way each year. However, the findings are similar whether we consider the upper or lower lines. In this report, households defined as poor are households below the upper poverty line, which categorizes more households as poor than does the less-inclusive lower poverty line.

In Table III.12 we present the impacts on household consumption and poverty. The estimated impact on total household consumption was positive but not statistically significant. However, there is a statistically significant estimated impact on poverty, with the treatment households being 10 percentage points more likely to be poor than the comparison households (p-value: 0.06). This non-negative (that is, detrimental) impact on poverty is concentrated among male-headed households, on which the impact on poverty is 13 percentage points and is significant at the 5 percent level (Table III.13). The impact estimate of an increase in poverty is at odds with the estimated (insignificant) increase in household consumption, whether measured in total (as reported) or per person (not shown); this may occur because the RRRP has different effects on households that were near the poverty line before the RRRP was implemented than for those well above the poverty line. (See Tables A.3a, A.3b, and A.3c in Appendix A for baseline and follow-up distributions of consumption.) Unfortunately, the cross-sectional nature of the ILCS precludes exploring this idea with additional analyses. The detrimental estimated impact on poverty is a result for which we do not have a compelling explanation, and may be an anomaly related to the mix of households sampled at baseline versus follow-up. As alternative measures of household well-being, we also considered savings and debt, neither of which exhibited evidence of impacts, corroborating our interpretation that there were not meaningful effects on household well-being in this timeframe.

As a sensitivity test, we also considered whether impacts are substantively different when we examine impacts on log income and log consumption instead of the levels for these variables. With this transformation, the estimated treatment effects provide the estimated impacts on consumption and income in percentage terms rather than levels. Because baseline economic outcomes are not equivalent for the treatment and comparison groups, it is possible that one group's economic outcomes improved faster in percentage terms than the other, even if the absolute level of the change was equivalent. Converting outcomes to log values also has the advantage that it suppresses the possible influence of large outlying values. These results are detailed in Appendix A. However, neither the estimated impacts on income net of remittances nor the estimated impacts on log economic outcomes substantively change our core findings.

Table III.12. Impact of road rehabilitation on household consumption and poverty (averages in US$ unless otherwise indicated)

	Treatment Mean	Comparison Mean	Program Impact	p-Value of Impact
Total household consumption	3,483	3,277	206	0.49
Poor	38	27	10*	0.06
Households	2,560	2,288		
Road Links	27	28		

Source: 2007 and 2011 ILCS.

Note: All dollars are in 2011 dollars, converted from Armenian Drams using the OANDA currency. Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place.

*/**/*** statistically significant at 10, 5, and 1 percent, respectively

Table III.13. Impact of road rehabilitation on household consumption and poverty, by gender of head of household (averages in US$ unless otherwise indicated)

	Overall Impact: All Households	Program Impact: Female-Headed Households	Program Impact: Male-Headed Households
Total household consumption	206	435	122
Poor	10*	4	13**
Households	4,731	1,346	3,385
Road Links	55	55	55

Source: 2007 and 2011 ILCS.

Note: All dollars are in 2011 dollars, converted from Armenian Drams using the OANDA currency. Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place.

*/**/*** statistically significant at 10, 5, and 1 percent, respectively

E. Implications for RRRP's economic rate of return

As described in Chapter I, both MCC and the World Bank used estimated ERRs to determine whether road projects would be funded. For MCC, this was a key criterion in selecting roads from among many that were considered. For the World Bank, ERRs played a less central role in selecting road projects, but only projects with ERRs high enough to justify the investment were financed. In both cases, MCC and the World Bank projected ERRs based on expected benefits using software (HDM-4) that takes road roughness and traffic counts into account.

The ultimate objective of the RRRP, however, is to increase the total household income of all program beneficiaries by a large enough amount that the program costs are justified. Thus, the ideal way to calculate ex post ERRs would be based on summing the impact of the RRRP on every beneficiary household for each year after rehabilitation began,[24] and using that stream of benefits to estimate the rate of return on the cost outlays. This calculation is unfortunately not feasible at present for two reasons. First, as described previously, the time frame for the present study is too short for us to expect the longer-term RRRP impacts to have fully materialized, including impacts on household income. As such, it is too early to determine if the benefits justify the costs. Second, the sampling strategy focuses on the benefits that accrue to households in communities that are directly connected by the roads that were linked, and does not consider households in other nearby communities who may nevertheless benefit (albeit to a lesser extent). A model based on this smaller sample of nearby communities would likely underestimate the full range of benefits experienced by all affected communities.

If ILCS data were to be available for the treatment and comparison road links in the future, it could be feasible to estimate RRRP impacts on longer-term income for communities directly served by rehabilitated roads. Using that annual income impact estimate, an ERR could be calculated based on the benefits that accrue only to communities directly served by project roads, which would be a lower bound on the ERR if all beneficiaries were considered. The calculations could also assume that the impacts on communities that are indirectly served by RRRP road links are as large as impacts for those directly served; with this assumption, an upper bound for the ERR could be calculated. With either of these approaches, however, an important caveat is that household income is highly variable and its impact estimates are likely to be measured imprecisely, even if the point estimate is large. Consequently, the estimated ERR will also be imprecise.

It is also possible to use the same HDM-4 approach to recalculate ex post ERRs that were used to estimate ex ante ERRs. Mathematica staff is currently working with the World Bank to try to obtain that information so that those ERRs could be presented in a future draft of this report.

[24] MCC policy on ERRs is to estimate benefits and costs over a 20-year horizon.

This page has been left blank for double-sided copying.

IV. CONCLUSIONS

A. Summary of findings

To summarize the results of this impact analysis, it is useful to revisit the program logic. Starting with immediate and short-term outcomes and proceeding to medium- and long-term outcomes, we can assess the evidence regarding the RRRP's impact on each outcome. As illustrated in Table IV.1, there is strong evidence of large impacts of the RRRP on immediate outcomes and short-term outcomes of improved road quality, less travel time, greater access to markets, and increased vehicular activity. In particular, the 40 percentage point increase in favorability ratings of regional roads among treatment households relative to comparison group households represents a dramatic improvement. However, we found no evidence of increased short-term employment income generated by rehabilitation efforts.

Given large impacts on several immediate and short-term outcomes, an examination of medium-term impacts is instructive to see if there is evidence of early impacts on these dimensions, even if we do not expect the full impacts to have materialized in the short follow-up time frame. In examining medium-term impacts, we hope to determine if improved road quality, market access, and transportation services actually changed household investment and production decisions. The most positive finding regarding investments is treatment households' apparent investments in sheep and other animals relative to the comparison group. However, it should be noted that a treatment-comparison difference of $31 in animal purchases over the course of a full year is small in magnitude. In addition, there is some scattered evidence of impacts in production—particularly production of jams and preserved vegetables—but those positive impacts are small, and there is also evidence of negative impacts on egg sales. Overall, these findings regarding production are more likely anomalies due to the many outcomes that were examined, as they are not part of a broader pattern of positive impacts on production nor is there a theoretical reason that these types of production would have been impacted more strongly. Given the paucity of impacts related to investment and production, it is unsurprising that the analysis revealed no impacts of the RRRP on sales in this time frame.

As we see no strong evidence of medium-term impacts on investment, production, and sales, it is also unsurprising that our analysis of long-term outcomes of household income, consumption, and poverty also revealed no impacts. As stated above, this is expected given the time frame of the evaluation—between one and two years following road rehabilitation. Particularly interesting is that more than one year after construction, treatment households appeared to alter their social and consumer behavior more than their agricultural and labor market behavior. For example, the analysis revealed that treatment households were 17 percentage points more likely to use roads for "other" purposes—particularly to visit relatives and shop—but no more likely to use roads to get to work. There is also no strong evidence that treatment group farmers altered the mix of crops they produced and sold, or that the prices they received for these crops changed as a result of road rehabilitation.

Table IV.1. Evidence assessment for impact of road rehabilitation investments

Outcomes in Program Logic		Evidence Assessment
Immediate	Improved road quality	Strong evidence of large impacts (39 percentage point increase in favorability rating of regional roads)
	Reduced vehicle operating costs	Strong indirect evidence of large impacts (19 percentage point increase in approval for transportation services; 17 percentage point increase in use of roads for noncommercial purposes; decrease in time spent using roads to sell agricultural production)
	Reduced travel time	
	Nonpermanent employment linked to construction	Inconclusive; main specification finds no evidence, but alternative specifications reveal possible positive impacts
	Improved access to markets	Strong evidence of large impacts in market access (20 percentage point decrease in market access difficulties)
Short term	Improved access to social infrastructure	No evidence of impacts
	Increased vehicular activity: commercial	Some evidence of impact (increase in use of roads to buy agricultural inputs and decrease in days roads were used to sell agricultural production)
	Increased vehicular activity: non-commercial	Strong evidence of large impacts (17 percentage point increase in use of roads for noncommercial purposes);
Medium term	Increased investment	Limited evidence of small impacts (1.5 more sheep owned by treatment households and $31 increase in annual animal purchases)
	Increased employment	No evidence of impacts
	Increased production	Limited evidence that may be anomalous (5 percentage point increase in jam production, 19 kg increase in preserved vegetable production, but $31 decrease in egg sales)
	Increased transactions	No evidence of impacts
Long term	Increased household income	No evidence of impacts
	Increased household consumption	No evidence of impacts
	Reduced rural poverty	Some evidence of increase in rural poverty; likely an anomaly due to sample composition

B. Limitations of the evaluation

Although we believe the evaluation findings that we report provide valuable evidence on the short-term impacts of the RRRP, there are several important limitations of this analysis that should be reiterated and borne in mind in drawing conclusions. First and perhaps most important, as is discussed in the preceding section, the follow-up time frame is short relative to the period during which the full impacts on longer-term outcomes such as household income would be fully realized. Previous work examining impacts of road rehabilitation in Vietnam (van de Walle and Mu 2007) found that impacts on both the longer-term outcomes as well as the pathways to those longer-term outcomes materialized by four years after rehabilitation; however, no such impacts were evident when van de Walle and Mu estimated impacts approximately two years after rehabilitation. Although their context is much different, it does suggest that longer-term impacts of the RRRP could manifest in future years, despite their apparent absence in the medium term.

The evaluation methodology also has two important limitations. First, in any non-experimental comparison group evaluation design such as ours, there may be systematic

differences between treatment and comparison communities—besides access to rehabilitated roads—that could affect key outcomes in the evaluation. To the extent that the comparison group differs from the treatment group along dimensions important for outcomes, we are unable to distinguish actual program impacts from underlying differences in the two groups unless we can credibly identify and control for these preexisting differences. Factors such as communities' political power, local leadership, social capital, and social empowerment all could bias impact estimates in comparison studies such as ours include political power, local leadership, social capital, and social empowerment. Additionally, the estimation strategy accounts for preexisting differences in baseline outcomes, but not preexisting differences in the *trends* of those outcomes. We unfortunately do not have sufficient pre-rehabilitation data to test whether the treatment and comparison communities had similar growth rates for the key outcomes, such as household income, poverty, and so on. That our findings are generally robust to alternative estimation strategies bolsters the interpretation that bias is less likely to explain our results, but we cannot conclude that our estimates are not biased to some extent by unobserved factors or different trends.

Finally, our estimates are not always statistically precise, because of features of the evaluation design. Instead of having households drawn independently from many different communities, they are drawn from just 56 road projects. This reduces statistical precision to do community-level clustering effects. Moreover, some of the key outcomes, such as annual household income, are highly variable, making it difficult to precisely estimate impacts even if our sample were large and unclustered.

C. Policy implications

It is difficult to assess the policy implications of these findings after the relatively short follow-up period of this evaluation, given that it is unclear if road rehabilitation efforts will eventually generate desired outcomes of increased agricultural productivity and household income in future years, especially when coupled with the imprecision of some of the impact estimates for key outcomes such as household income. Thus, we can do little but speculate on the policy implications. However, one sound implication of these findings is that road rehabilitation efforts appear capable of altering households' activities and use of transportation services in the short term, but they probably were not sufficient to stimulate agricultural production and sales in a time frame of one to two years.

In particular, the relatively long six-month agricultural cycles in Armenia may lead to relatively slow changes in agricultural investment and production that are unlikely to be detected in one year, or even two years. Although immediate improvements in market access could help a farmer obtain cheaper inputs or sell his current production for a higher price, the same farmer cannot change his mix of crops or agricultural products in response to improved market access until the subsequent agricultural season. Often, changing crops and agricultural products is a process of trial and error, and of incremental increases in investment, production, and income. As such, a full transition to a new high-value crop or a new processed food would likely take several years to generate measurable increases in sales and income. It bears mention, however, that there is also no evidence that prices farmers receive for their products have changed either, even though prices could have changed more rapidly than agricultural investments.

D. Dissemination plans

Mathematica staff will travel to Armenia in 2015 to present the findings of this RRRP impact analysis as well as findings from the irrigation impact evaluation to key stakeholders, particularly World Bank staff and Armenian government officials who are involved in rural infrastructure activities.

REFERENCES

Armenia Roads Directorate. 2008. *Monitoring and Evaluation Report.* Yerevan, Armenia: Armenia Roads Directorate.

Banerjee, A. and E. Duflo. 2009. "The Experimental Approach to Development Economics." *Annual Review of Economics,* Annual Reviews, 1(1): 151-178.

BenYishay, A. and R. Tunstall. 2010. "Comparing Evaluations of Road Improvements: An Organizational View." Working paper.

Deaton, A. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* Washington, DC: World Bank.

Dercon, S., D. Gilligan, J. Hoddinott, and T. Woldehanna. 2008. "The Impact of Agricultural Extension and Roads on Poverty and Consumption Growth in Fifteen Ethiopian Villages." IFPRI Discussion Paper 00840.

Escobal, J. and C. Ponce. 2002. "The Benefits of Rural Roads: Enhancing Income Opportunities for the Rural Poor." Grupo de Análisis para el Desarrollo Working Paper 40-I.

Estache, A. 2010. "A survey of impact evalutions of infrastructure projects, programs, and policies." ECARES Working Paper 2010-005, Université Libre de Bruxelles.

Fortson, K., P. Gleason, E. Kopa, and N. Verbitsky-Savitz. 2013a. "Horseshoes, Hand Grenades, and Treatment Effects? Reassessing Bias in Nonexperimental Estimators." Working paper. Oakland, CA: Mathematica Policy Research, March 2013

Fortson, K., A. Rangarajan, R. Blair, J. Lee, and V. Gilbert. 2013b. "Evaluation of Water-to-Market Training in Armenia." Final report submitted to the Millennium Challenge Corporation.

Gannon, C. and Z. Liu. 1997. "Poverty and Transport." World Bank TWU-30.

Gibson, J. and S. Rozelle. 2003. "Poverty and Access to Roads in Papua New Guinea," *Economic Development and Cultural Change* 52, pp. 159-185.

Gonzalez-Navarro, M. and C. Quintana-Domeque. 2012. "Paving Streets for the Poor: Experimental Analysis of Infrastructure Effects." Working paper.

Imbens, G. 2010. "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)," *Journal of Economic Literature,* 48(2): 339-423.

Imbens, G., and J. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature,* vol. 47, no. 1, pp. 5–82.

Khandker, S., Z. Bakht, G. Koolwal. 2006. "The Poverty Impact of Rural Roads: Evidence from Bangladesh." World Bank Policy Research Working Paper 3875.

Lokshin, M. and R. Yemtsov. 2005. "Has Rural Infrastructure Rehabilitation in Georgia Helped the Poor?" *World Bank Economic Review* 19(2), pp. 311-333.

Casaburi, L., R. Glennester, and T. Suri. 2012. "Rural Roads and Intermediated Trade: Regression Discontinuity Evidence from Sierra Leone." Working paper.

MCA-Armenia. 2011. *The Program is Over: All About Results.* Yerevan, Armenia: Millennium Challenge Account-Armenia.

MCC. 2007. *Millennium Challenge Compact with Armenia, Executive Summary.* Washington, DC: Millennium Challenge Corporation.

MCC. 2011. *Armenia Compact: Economic Rates of Return.* http://www.mcc.gov/pages/countries/err/armenia-compact, accessed August 1, 2013.

MCC. 2012. *Table of Key Performance Indicators (Closeout Results: October to December 2011)*. Washington, DC: Millennium Challenge Corporation.

Mu, R. and D. van de Walle. 2007. "Rural Roads and Local Market Development in Vietnam." World Bank Policy Research Working Paper 4340.

National Statistical Service. 2010. *Measuring Poverty in Armenia: Methodological Explanations.* Yerevan, Armenia: National Statistical Service of Armenia.

Van de Walle, D. 2009. "Impact Evaluation of Rural Road Projects," *Journal of Development Effectiveness* 1(1), pp. 15-36.

World Bank. 2010. *Transport Notes Improving Local Roads and Creating Jobs through Rapid Response Projects: Lessons from Armenia Lifeline Roads Improvement Project.* Washington, DC: World Bank.

World Bank. 2012. *Project Appraisal Document on a proposed loan in the amount of $45 million to the Republic of Armenia for the Lifeline Road Network Improvement Project.* Washington, DC: World Bank.

World Bank. 2013. *Implementation Status & Results Armenia Lifeline Roads Improvement Project (P115486).* Washington, DC: World Bank.

World Bank. June 2014. *Implementation Completion and Results Report: Lifeline Roads Improvement Project (IDA-45490 IBRD-77510 IBRD-79360 TF-94897).* Washington, DC: World Bank.

APPENDIX A

SENSITIVITY ANALYSES AND SUPPLEMENTAL CALCULATIONS

This page has been left blank for double-sided copying.

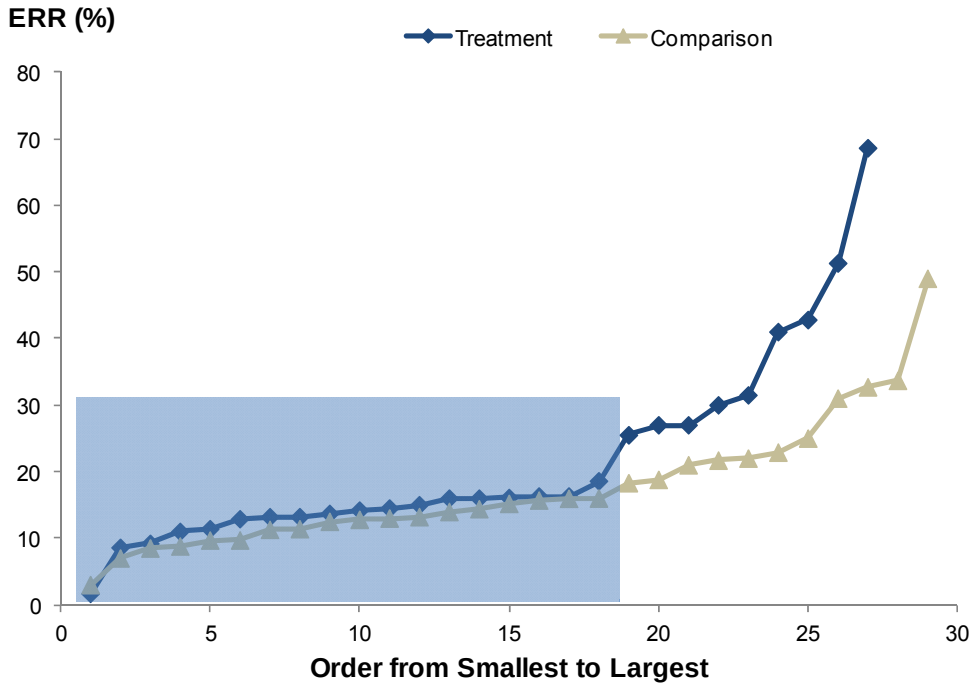
We conducted several sensitivity tests to examine the extent to which the key impact estimates are affected by alternative empirical specifications. As described in Chapter I, non-experimental evaluation designs assume that all factors that are correlated with both treatment status and the key outcomes can be adequately accounted for in the empirical specification. Different non-experimental methods account for these factors in different ways and have different theoretical advantages. Even if all of the factors that would potentially bias the RRRP impact estimates are present in the ILCS data, the presence of significant baseline differences between our treatment and comparison groups makes it crucial that the model appropriately controls for these factors in the regression models. The six sensitivity analyses that we conducted vary along three dimensions: (1) Whether regression or weighting is used to account for observable differences, (2) Which observable characteristics are accounted for, and (3) Restrictions on the road links included in the analysis.

In the remainder of this appendix, we describe each of the six alternative specifications that we examined and how their impact estimates compare with those of the main specification. Taken as a whole, the impact estimates on key outcomes are robust to the alternative specifications we considered—more so than we expected at the outset of this analysis.

The first alternative specifications consider the possibility that, with a relatively small number of road links in the analysis (55), controlling for year by ERR interactions as well as region of the country (the 10 marzes of Armenia, excluding the capital city of Yerevan) could generate instability in the model. We thus consider more parsimonious specifications. First, we estimated a simple difference-in-difference model in which we did not control for either marz or ex ante ERRs at all **(Alternative 1)**. Second, we also examined the extent to which controls for the interaction between ERR and year are altering estimates **(Alternative 2)**, as opposed to simply controlling for ERR. As described in Chapter I, the main specification was structured to allow for the possibility that ERRs could indicate differential changes in key outcomes for the treatment and comparison groups over time. This was partially motivated by the original regression discontinuity evaluation design. However, given that we found that ex ante ERRs were only weakly correlated with treatment status, there is less of an empirical justification for featuring the ERRs as prominently in the regression model.

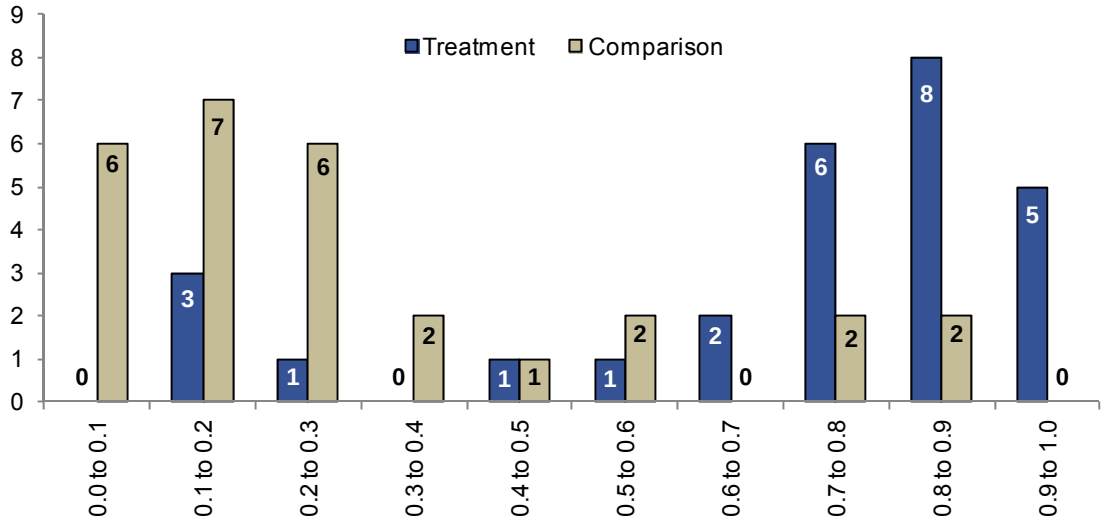
A related observation is that, as highlighted in Figure A.1, the smallest 18 ex ante ERRs among the treatment road links are very similar to the smallest 18 ex ante ERRs among the comparison group. The larger ERRs diverge more for the two groups. Thus, for this subset of 36 road links, the ERR did not determine treatment status at all. We examined the extent to which restricting to those roads with similar ERRs altered impact estimates in **Alternative 3**. This specification has the advantage that it is less reliant on correctly modeling the functional form of the ERRs in the model to account for treatment-comparison differences in the ex ante ERRs, possibly reducing bias. However, this smaller set of road links leads to weaker statistical precision, and the impact estimates apply only to the subset of treatment roads that have smaller ERRs.

Figure A.1. Estimated ex ante ERRs for treatment and comparison road links



Another set of sensitivity analyses use propensity score methods to reweight the treatment and comparison road links such that they are more similar to each other on observable baseline characteristics. (See Imbens and Wooldridge 2009 for a detailed discussion on propensity score methods.) We used community-level averages of baseline economic and demographic characteristics as well as the ERRs to predict the probability that a given road link was in the treatment group; this predicted probability is the estimated propensity score. In **Alternative 4,** we then weighted treatment road links by the inverse of their propensity scores and weighted comparison road links by the inverse of 1 minus their propensity scores. Conceptually, treatment roads that were less likely to be treated get greater weight, and comparison roads that were more likely to be treated get greater weight. This approach has the advantage that it potentially reduces some bias that may have occurred due to differences in the observable characteristics of treatment and comparison roads. However, it is also susceptible to very small or large estimated propensity scores that can inflate standard errors by creating highly variable weights. As illustrated in Figure A.2, the treatment group does not have many roads that were unlikely to have been selected based on their observable characteristics, nor does the comparison group have any roads that were likely to be selected. This imbalance motivates our final sensitivity analysis, **Alternative 5,** in which we used the same propensity score approach but removed roads in either the treatment or comparison group that were poorly matched to the roads in the other group—those roads that did not share a common probability support. Specifically, we removed from the analysis sample those roads with propensity scores of less than 0.1 or greater than 0.9 because there was not at least one treatment link and one comparison link at these bands of the score's distribution (see the left and right ends of Figure A.2). Although this resulted in a more evenly matched analysis sample, it also reduced sample size and precision in our estimates, and as with Alternative 3, the impact estimates only generalize to a subset of the roads.

Figure A.2. Estimated propensity scores for treatment and comparison road links



Impact estimates for a full and proper propensity score analysis would account for estimation error in the propensity score model itself as well as the regression model (Imbens and Wooldridge 2009), usually using bootstrapping. Correctly accounting for estimation error in the propensity score model does not affect the estimated impacts but can reduce statistical precision, though in practice it does not usually make a meaningful difference.[25]

Lastly, to explore whether the roads that were rehabilitated in 2010 dampen the results because of the short period between their completion and data collection in 2011, **Alternative 6** restricts the analysis to treatment roads that were rehabilitated in 2009 (and retains all comparison roads). Table A.1 summarizes the six alternative sensitivity tests and the rationale for each.

[25] Because this is a sensitivity check, we do not account for estimation error in the propensity score model. Thus, the estimated impacts that we report for Alternatives 4 and 5 may be more precise than they should be.

Table A.1. Description of sensitivity tests, roads analysis

Sensitivity Test	Rationale
Unadjusted difference-in-difference (Alternative 1)	Examine the extent to which controlling for ERRs and region is altering estimates or reducing statistical precision due to the loss of degrees of freedom.
Exclude ERR x Year interaction controls (Alternative 2)	Determine the extent to which controls for the interaction between ERR and year are altering estimates. Specifically, we remove the ERR x Year interaction terms, and instead include a single ERR term in the model.
Trimming the sample by ERR (Alternative 3)	Examine if keeping only those roads in either the treatment or comparison group with an ERR similar to that of a road in the other group results in substantially different impact estimates. Although this results in a more evenly matched analysis sample, it also reduces sample size.
Weighting by propensity scores (Alternative 4)	Examine if using propensity scores to reweight the sample as if there were a similar distribution of baseline characteristics between the treatment and comparison groups that results in substantially different impact estimates.
Trimming the sample by propensity scores (Alternative 5)	Similar to Alternative 4 but also exploring whether restricting the analysis to the treatment and comparison road links that are most similar to each other affects the estimates.
Restricting the sample to roads rehabilitated in 2009 (Alternative 6)	Examine if focusing on roads that were rehabilitated earlier reveals impacts that are obscured because of roads with only one year between rehabilitation and follow-up.

As shown in Table A.2, these sensitivity tests revealed that the impacts of road rehabilitation efforts on key intermediate outcome measures are robust to various transformations and alternate specifications. The 39 percentage point impact in our main specification ($p < 0.01$) on the likelihood of households rating local roads as "good" or "excellent" was supported across all alternative specifications. All estimated impacts were positive with a magnitude of at least 36 percentage points, and statistically significant at the 1 percent level. The positive impact on the likelihood of households reporting no problems with market access (20 percentage points; $p < 0.01$) was also robust to all alternative specifications. All estimated impacts had magnitude of at least 15 percentage points, and in all five sensitivity tests estimated impacts were statistically significant at the 5 percent level or less. The estimated impact on the likelihood of households using roads for a purpose other than buying agricultural inputs, selling agricultural products, or going to work (17 percentage points in our main specification, $p = 0.01$) was also robust to alternative specifications. Nearly identical impacts were estimated using Alternative 1 and Alternative 2, and all were significant at the 5 percent level. We estimated a somewhat larger impact using propensity scores to balance treatment and comparison groups (28 percentage points), trimming the sample using ERRs (27 percentage points), and weighting and trimming the sample by propensity scores (30 percentage points).

Table A.2. Impact estimates for key outcomes, by alternative estimation techniques

	Main model	Alternative 1 (simple diff-in-diff)	Alternative 2 (no ERR x Year interactions	Alternative 3 (trimming the sample by ERR)	Alternative 4 (weighting by propensity scores)	Alternative 5 (weighting and trimming the sample by propensity scores)	Alternative 6 (restricting treatment to 2009 rehabilitation)
Regional roads good or excellent (%)	39***	36***	36***	37***	40***	45***	39***
Transportation services good or excellent (%)	19*	19	20**	19*	31**	34**	18*
No problems with market access (%)	20***	21***	21***	28***	16***	15**	22***
Use roads for other purposes (%)	17**	16**	16**	27***	28***	30***	17**
Total expenditure on animal purchases ($)	31*	28*	28*	40**	40	43	36
Total income from crop sales ($)	-99	-159	-160	-90	172	222	-173
Total income from food sales ($)	12	-7	-1	-102	-59	-118	-10
Total household income ($)	-431	-560	-567	-522	85	121	-665
Employment income ($)	-200	-296	-294	-403	168	229	-123
Log household income	-0.03	--	-0.07	-0.10	0.05	0.05	0.02
Income from other sources ($)	72	58	54	89	130**	140**	64
Non-remittance income ($)	-82	-221	-221	-471	344	322	-127
Household consumption	206	261	259	243	485	558	59
Log household consumption	0.07	--	0.09	0.06	0.12	0.14	0.05
Poverty rate (%)	10*	8	8	13*	4	1	12*
Households	4,848	4,848	4,848	2,462	4,848	3,976	3,440
Road Links	55	55	55	36	55	44	45

Note: All dollars are in 2011 dollars, converted from Armenian Drams using the OANDA currency. Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place.

*/**/*** statistically significant at 10, 5, and 1 percent, respectively

We also estimated a positive impact on the likelihood of households rating transportation services as "good" or "excellent" using the main specification (19 percentage points, p = 0.05). All but one alternative estimation yielded statistically significant positive estimated impacts. Three of the five alternative specifications were significant at the 5 percent level, and one was significant at the 10 percent level.

Our main analysis found no significant impacts on total income from crop sales or total income from food sales, and these results are robust to alternative specifications. A statistically significant impact on total expenditure on animal purchase was estimated in our main analysis ($31, p = 0.06). Three of five alternative specifications yielded significant impacts of similar magnitude, although only the impact estimated after trimming the sample by ERR was significant at the 5 percent level.

Our main analysis also found no statistically significant impacts on economic outcomes, and these results are robust to alternative specifications. The estimated impacts on total household income, employment income, and household consumption were statistically insignificant across all specifications, although generally equal to our main findings in size and direction. We observed more positive impacts on economic outcomes using propensity score weights. The estimated impact on income from other sources using propensity score weights was statistically significant with and without trimming ($140 and $130, respectively); estimated impacts were insignificant using all other specifications. We estimated a positive, marginally significant impact on household poverty status in our main analysis (10 percentage points, p = 0.06). However, we found no significant impacts across four of five alternative specifications, though the magnitudes of the estimates are similar to our main specification. Only the impact estimated after trimming the sample by ERR was marginally significant (13 percentage points, p = 0.09).

As an additional sensitivity tests for the estimated impacts on economic outcomes, we also examined impacts on other definitions of the economic outcome measures. First, we examine impacts on income net of income from remittances. Income from remittances does not represent an improvement in productivity, only a transfer, so looking at income net of remittances tells us about possible program impacts on income-generation. Second, we consider whether impacts are substantively different when we examine impacts on log income and log consumption instead of the levels for these variables. With this transformation, the estimated treatment effects provide the estimated impacts on consumption and income in percentage terms rather than levels. Because baseline economic outcomes are not equivalent for the treatment and comparison groups, it is possible that one group's economic outcomes improved faster in percentage terms than the other, even if the absolute level of the change was equivalent. For example, as reported in Table II.5, the treatment group's household income in 2007 was $2,439 at baseline compared with $2,869 for the comparison group. If income increased by $500 for both groups, this would be an increase of about 21 percent for the treatment group but only 17 percent for the comparison group. Converting outcomes to log values also has the advantage that it suppresses the possible influence of large outlying values. However, neither the estimated impacts on income net of remittances nor the estimated impacts on log economic outcomes substantively change our core findings.

In contrast, as shown in Table A.3, these same sensitivity tests showed that our finding that road rehabilitation had no significant impact on short-term employment income in the year of rehabilitation ($2, p = 0.99) was not robust to alternative specifications. When we trim the sample by the ERR (Alternative 3), the estimated impact is large and positive ($357) and statistically significant at the 10 percent level (p = 0.09). The estimate is not significant in any other specification, though the magnitudes of the estimated impacts using Alternatives 4 and 5 ($216 and $257, respectively) are similar to the estimate for Alternative 3. Based on the weight of the evidence, the null finding from the main specification remains our best estimate of the impact on short-term impact, but this finding is not conclusive in light of its lack of robustness.

Table A.3. Impact estimates for short-term employment, by alternative estimation techniques

	Main model	Alternative 1 (simple diff-in-diff)	Alternative 2 (no ERR x Year interactions)	Alternative 3 (trimming the sample by ERR)	Alternative 4 (weighting by propensity scores)	Alternative 5 (weighting and trimming the sample by propensity scores)
Employment in year of rehab	2	-119	-103	357*	216	257

Note: We do not test Alternative 6 for this outcome because this is exclusively a short-term outcome, and Alternative 6 is designed to examine a longer time horizon. All dollars are in 2011 dollars, converted from Armenian Drams using the OANDA currency. Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place.

*/**/*** statistically significant at 10, 5, and 1 percent, respectively

Impacts by land holdings. To test whether road improvements affected larger land owners differently from smaller land owners, we estimated the impact of the RRRP separately for households with one or more hectare of land versus households with less than one hectare of land. Estimating impacts on households separately by land ownership shows only small differences in outcomes between households with less than one hectare of land and households with at least one hectare of land. The impact on the share of households who rated transportation services as good or excellent was positive on both groups, but only statistically significant in households with at least one hectare of land (27 percentage points, p = 0.02). The estimated impacts on total animal expenditures was also only statistically significant in households with at least one hectare of land ($55, p = 0.04).

However, the impacts on the main outcome measures are consistent with those observed in the main analysis. Impact estimates on total household income were negative but statistically insignificant for both groups. The impacts on the share of households who rated regional roads as good and reported problems with market access were statistically significant and similar in magnitude and direction in both groups as the impacts estimated in the main analysis.

Households with less than one hectare of land appear to be driving the positive and statistically significant (at the 10 percent level) impact observed on household consumption ($624, p = 0.07). The estimated impact on total consumption of households with at least one hectares of land was negative, but statistically insignificant (-$80, p = 0.84). The impacts on income and consumption are qualitatively similar when we examine impacts on log income and log consumption.

Table A.4. Estimated Impacts, by household land ownership

	Households with 0-1 hectares of land	Households with 1+ hectares of land
Regional roads good or excellent (%)	43***	35***
Transportation services good or excellent (%)	6	27**
Problems with market access (%)	-20**	-17***
Use roads for other purposes (%)	17**	15
Total expenditure on animal purchases	-2	55**
Total income from crop sales	-212	4
Total income from food sales	-33	70
Total household income ($)	-520	-379
Log total household income	-0.03	-0.04
Employment income ($)	-24	-309
Income from other sources ($)	86	50
Non-remittance income ($)	-6	-179
Household consumption	624*	-80
Log household consumption	0.18**	-0.02
Poverty rate (%)	9	13*
Employment income in year of rehabilitation	-62	-23
Households	2,615	2,273
Road Links	55	55

Note: All dollars are in 2011 dollars, converted from Armenian Drams using the OANDA currency. Impact estimates are the average difference between treatment and comparison households for a given outcome measure, controlling for baseline differences between the two groups as described in Chapter I. Cases in which program impact is not the difference between treatment and comparison means reflect rounding to the nearest percentage point or decimal place.

*/**/*** statistically significant at 10, 5, and 1 percent, respectively

Consumption distributions. To provide additional context for impact estimates on household consumption and poverty—which is based on consumption per adult equivalent—we illustrate the distribution of household consumption in Figures A.3a, A.3b and A.3c. As shown in the first two figures, the distribution of household consumption was balanced at baseline. In both 2007 and 2008, the difference in consumption among treatment and comparison households is less than $200 at nearly all percentiles. An exception is the 95th percentile in both years—the portion of households with consumption that was higher than 95 percent of all households in the sample—for which there is a difference between the treatment and comparison group of at least $400. At baseline, comparison households had slightly higher consumption than treatment households at nearly all percentiles. However, the distribution shifted at follow-up, such that treatment households had higher consumption at all percentiles. Notably, the distribution of consumption for comparison households shifted in 2011—consumption at all percentiles was lower in 2011 than in either baseline year. At follow-up there are more prominent differences on the order of $400 for the 25th-75th percentiles, which is larger than the estimated average impact of $236, but remains statistically insignificant.

Figure A.3a. Distribution of total household consumption, 2007

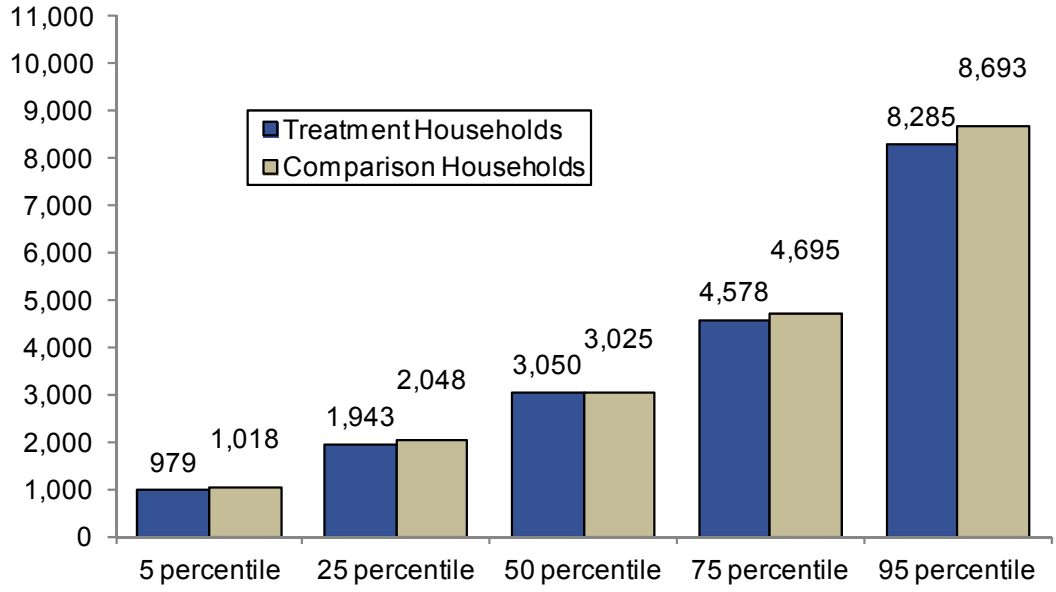


Figure A.3b. Distribution of total household consumption, 2008

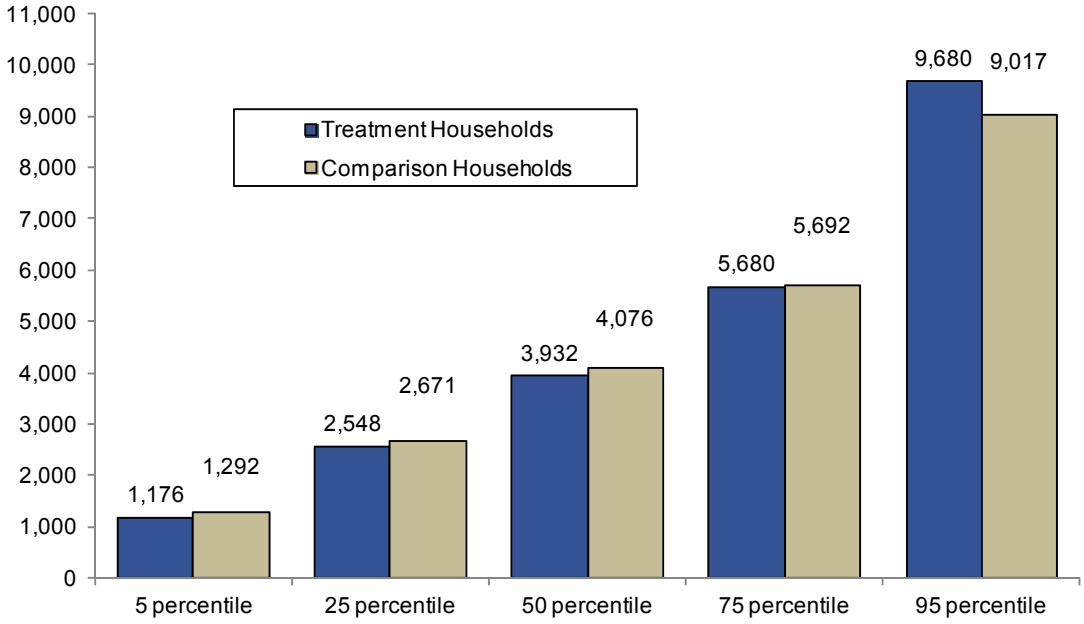
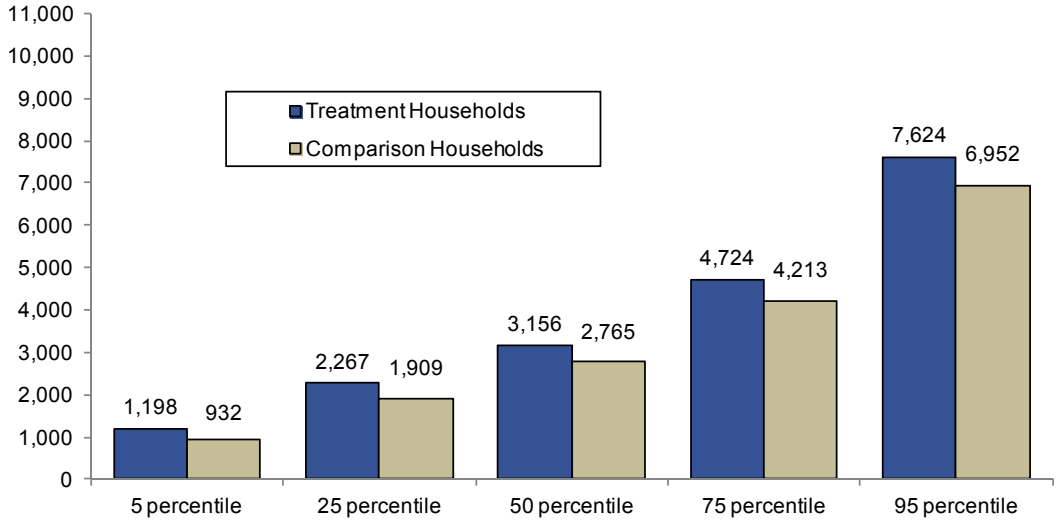


Figure A.3c. Distribution of total household consumption, 2011



Documenting migration patterns. We also considered the possibility that migration patterns could lead to misleading impact estimates of the program. For example, if rehabilitated roads attracted new residents, this could change the socioeconomic composition of the communities, and these compositional changes could obscure impacts on non-migrating communities. To explore this possibility, we attempted to document recent migration patterns among households in the study sample for each year of the survey. The share of households who reported changing their dwelling place within the past five years was low at baseline for both treatment and comparison households (2 percent and 4 percent, respectively). In 2011, the difference between the share of treatment and comparison households that reported moving was less than one percent, and a smaller share of households in both groups reported moving within the last five years relative to the share at baseline (1 percent and 2 percent, respectively). Therefore, it is unlikely that migration patterns have a discernible influence on impact estimates.

Table A.5. Households that changed dwelling, within last 5 years

Year	Treatment	Comparison
	Percentage of Households	
2007	2.3	3.5
2008	2.4	3.5
2011	1.4	2.1

APPENDIX B

DETAILED REGRESSION RESULTS

This page has been left blank for double-sided copying.

Chapter I summarized the regression framework used to estimate impacts, and in this appendix we provide more detailed regression output for the estimates from key outcome measures. To recap, impacts were estimated using the following regression model:

$$(1) \quad y_{irt} = \beta' x_{ir} + \theta' R_t + \lambda T_r + \gamma R_{t=2011} T_r + \delta' R_t \times ERR_r + \eta_{rt} + \varepsilon_{irt}$$

where y_{irt} is the outcome of interest for household *i* served by road link *r* at time *t;* x_{ir} is a vector of time-invariant characteristics of household *i* served by road link *r* as well as road link–specific characteristics; T_r is an indicator equal to 1 if road link *r* is in the treatment group and 0 if it is in the comparison group; $R_{t\neg}$ is a vector of binary variables for each round of data included in the analysis; ERR is the economic rate of return associated with each road link; η_{rt} is a road-specific error term; ε_{irt} is a random error term for household i served by road link *r* observed at time *t;* and β, θ, λ, γ, and δ are parameters to be estimated.

We used households as the unit of analysis but clustered by road link. Our model also accounted for the specific nature of our data—repeated cross-sectional data, rather than panel data of the same households over time. The specific empirical framework we employed was difference-in-differences estimation, in order to estimate how outcomes changed for communities served by treatment roads before and after road rehabilitation (the first difference) compared to how outcomes changed for communities served by comparison roads over the same time period (the second difference). Household-level final weights were constructed from link weights—provided in the survey data to account for the population size of each survey community and the number of surveys conducted for each road link. Final weights for each household-level observation were the link weight divided by the sum of link weights for households in communities associated with the link, such that each link had the same total weight in the analysis.

Included in the right-hand side of the model are specific household demographic characteristics, series of binary variables for each year of the survey, a series of binary variables indicating geographic location (marz), a binary variable indicating treatment status, an interaction term between treatment status and the binary variable for 2011, and a series of interaction terms between economic rate of return (ERR) and survey year. The estimated ERR is interacted with the set of dummy variables indicating each year of the survey to control for the possibility that outcomes are not uniformly higher or lower in all years depending on ERR—for example, that all roads with high ERRs would have had higher post-intervention outcomes even if they were not treated.

Table B.1 presents the parameter estimates for all variables included in the model from the estimated regression models for three key outcomes: total household income, household consumption, and poverty status. The coefficient on the interaction of treatment status and the year 2011 indicator variable is the estimated impact of the RRRP on household income, consumption, and poverty.

Table B.1 Complete regression results for selected outcome measures

	Total household income		Household consumption		Poverty rate	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Female-headed household	-3.045	0.004	-2.705	0.946	0.044	0.017
Head of household age, 2007	15.875	0.000	-2.735	3.141	-0.062	0.067
Size of household	391.282	0.000	574.894	31.480	1.911	0.487
Head of household education level – less than secondary	-9.728	0.016	0.219	2.571	0.045	0.052
Head of household education level – full	-4.786	0.199	1.603	2.605	0.032	0.052
Head of household education level – secondary vocational	--	--	--	--	--	--
Head of household education level – higher than secondary *	9.134	0.035	10.509	2.765	-0.025	0.054
In 2007 *	--	--	--	--	--	--
In 2008	829.316	0.002	811.640	328.690	-0.953	4.986
In 2011	1748.181	0.000	-207.656	325.593	14.098	6.107
Aragatsotn	--	--	--	--	--	--
Ararat	-1447.201	0.006	225.348	439.180	10.430	6.203
Armavir	-1264.835	0.011	225.348	439.180	9.692	5.474
Gegharkunik	-478.557	0.347	-413.009	233.188	11.795	5.130
Lori	-1436.421	0.001	-60.349	245.907	18.273	4.725
Kotayk	-1102.120	0.012	671.567	224.792	0.753	4.271
Shirak	-1474.731	0.001	-286.844	230.869	18.237	6.086
Sjunik	-823.991	0.078	1121.058	319.418	3.677	4.382
Vayots Dzor	-2088.885	0.000	1147.844	285.337	8.405	4.422
Tavush	-1563.191	0.000	424.086	217.023	15.555	4.759
ERR x 2007	3.212	0.657	-7.990	6.639	-0.058	0.104
ERR x 2008	15.487	0.024	-1.673	7.364	-0.092	0.122
ERR x 2011	-8.617	0.438	2.828	6.421	-0.354	0.158
Treatment	-85.946	0.722	-194.788	218.247	-6.029	3.195
Treatment x 2011	-430.998	0.285	205.960	297.827	10.479	5.493

* Omitted category

To illustrate statistical precision of the impacts, we calculated minimum detectable impacts—the smallest true impacts that can be reliably detected—based on the standard errors from our analyses. Table B.2 presents minimum detectable impacts for many outcome measures. The minimum detectable impact is calculated as the standard error for the impact estimate for the specified outcomes times 2.80, which is the factor associated with a two-tailed t-test with 80 percent power and a 5 percent significance level. As discussed in Chapter II, the minimum detectable impacts are large mostly because of the limited number of road links that could be included in the analysis, which reduces statistical precision.

Table B.2. Impact estimate standard errors and minimum detectable impacts for key outcomes

Outcome Measure	Standard Error for Impact Estimate	Minimum Detectable Impact (Standard Error Multiplied by 2.80)
Percentage of households that rated regional roads as good or excellent	8.7	24.3
Percentage of households reporting problems with market access	5.5	15.5
Share of households that reported using a car	8.8	24.5
Share of households that reported using a bus	8.3	23.2
Shore of households that reported using a car or bus to access hospital	10.3	28.8
Share of household reporting distance to hospital less than 1 km	8.1	22.7
Distance to hospital, minutes	4.4	12.4
Percentage of households that used roads to purchase agricultural supplies	5.4	15.2
Percentage of households that used roads to access employment outside the community	2.7	7.5
Total cost of agricultural inputs, US$	55	154
Total cost of animals purchased, US$	17	48
Total value of crops sold, US$	186	521
Total value of animals sold, US$	88	247
Total household income, US$	402	1,125
Total household consumption, US$	298	833
Share of households that are poor according to their annual consumption	5.5	15.4

www.mathematica-mpr.com

Improving public well-being by conducting high quality, objective research and surveys

PRINCETON, NJ ■ ANN ARBOR, MI ■ CAMBRIDGE, MA ■ CHICAGO, IL ■ OAKLAND, CA ■ WASHINGTON, DC

MATHEMATICA
Policy Research