



Partnership for **FINANCE**  
in a **DIGITAL AFRICA**

# Approaches to determining the impact of digital finance programs



## ACKNOWLEDGEMENTS

Niamh Barry wrote this Snapshot with input from Jonathan Donner. This Snapshot was supported by the Mastercard Foundation.

---

## NOTES

The views presented in this paper are those of the author(s) and the Partnership, and do not necessarily represent the views of the Mastercard Foundation or Caribou Digital.

For questions or comments please contact us at [ideas@financedigitalafrica.org](mailto:ideas@financedigitalafrica.org).

---

## RECOMMENDED CITATION

Partnership for Finance in a Digital Africa, "Snapshot 15: Approaches to determining the impact of digital finance programs," Farnham, Surrey, United Kingdom: Caribou Digital Publishing, 2018.  
<http://www.financedigitalafrica.org/snapshots/15/2017/>.

---

## ABOUT THE PARTNERSHIP

The Mastercard Foundation Partnership for Finance in a Digital Africa (the "Partnership"), an initiative of the Foundation's Financial Inclusion Program, catalyzes knowledge and insights to promote meaningful financial inclusion in an increasingly digital world. Led and hosted by Caribou Digital, the Partnership works closely with leading organizations and companies across the digital finance space. By aggregating and synthesizing knowledge, conducting research to address key gaps, and identifying implications for the diverse actors working in the space, the Partnership strives to inform decisions with facts, and to accelerate meaningful financial inclusion for people across sub-Saharan Africa.

[www.financedigitalafrica.org](http://www.financedigitalafrica.org)

[www.mastercardfdn.org/financial-inclusion-program/](http://www.mastercardfdn.org/financial-inclusion-program/)

[cariboudigital.net](http://cariboudigital.net)



This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>.

Readers are encouraged to reproduce material from the Partnership for Finance in a Digital Africa for their own publications, as long as they are not being sold commercially. We request due acknowledgment, and, if possible, a copy of the publication. For online use, we ask readers to link to the original resource on the [www.financedigitalafrica.org](http://www.financedigitalafrica.org) website.

# Introduction

---

How do we know what's working? The number of digital finance products is growing rapidly, yet our knowledge around the client level effects of these products does not reflect this growth or diversity.<sup>1</sup> Moreover, it is difficult to find relevant comparisons and run experiments for products and services as complex as those of digital finance. When the most trusted and rigorous tests, such as Randomized Control Trials (RCTs), are not a viable option, how can we gather more impact insights without compromising quality? This Snapshot is for anyone running a digital finance program, many of whom might soon realize how much they have to contribute to the broader conversation about “what works.”

---

---

# There are “impact conversations”

At FiDA we think it is helpful to see impact assessment as a process more akin to a conversation than to the rendering of a verdict. One need not (and should not) rely on “capstone” documents shared only at the end of a successful project. Instead, impact conversations are broad, ongoing, multi-voiced dialogues about what is and is not working.

Different study designs inform impact conversations in different ways and at different stages. Impact conversations may begin as small discussions based on findings from exploratory qualitative and mixed method approaches. The discussion widens as contributors gather more information about product engagement from action research, administrative data, and rapid fire tests. Then, panels, pre-post tests and RCTs extend and deepen the conversation as insights on medium and longer-term effects are observed. Eventually, the conversation draws on the insights of diverse actors across countries, products, and clients.

The digital finance community must continue to push the industry forward with effective digital finance products and services. Rather than waiting for the results of long-term impact studies, we need to assess insights on impacts as they become available. Individually, a single impact insight may appear insignificant, but collectively they are powerful.

In this Snapshot, and in the spirit of encouraging conversation, we aim to lift the veil on different approaches to client impact measurement in order to expand the digital finance impact community and encourage the inclusion of various methods to advance the state of knowledge about the impacts of digital finance.

# We need you to be a part of these conversations

## Impact measurement matters

Broadly speaking, our community of practice is confident that the digitization of financial services presents an opportunity to promote financial inclusion and improve the well-being and livelihoods of low-income people around the world. But “flying blind” is not an option. Evidence about the implied causal pathways—that is, the impact of digital financial services on low-income households—is required for two reasons. First, in order to access philanthropic, social, and impact resources, the digital finance community must show evidence of digital finance’s effectiveness. Second, within the space, evidence about impact can guide programs such that they administer programs that effectively target certain populations and improve the efficacy of investments.

## Digital finance is hyped and heterogenous

While digitization presents a new avenue to contribute to financial inclusion, hype about design and delivery features remains. Significantly, the allure of trends such as big data, machine learning, and blockchain, may lead to “products first, problems second” scenarios. In reality, we have limited evidence of the impact of second generation digital finance products and services.<sup>2</sup>

Digital finance programs are often comprised of numerous components, have multiple implementing partners, and behave differently in different markets and with different client segments. Consequently, the nature of a digital finance program affects the viability of applying experimental designs to determine impact. There are several distinctions within digital finance programs:

- **Agile nature of product design and delivery:** From launch, products may change considerably as digital finance providers fix bugs or pivot based on market responses. In this dynamic scenario, an experimental design is not ideal.
- **Business growth:** Research designs that require control groups may not be an option for businesses; that is, withholding a product via a control group may negatively affect business growth or allow competitors to move in.
- **Market generalizability:** Designs that use a control group need to minimize attrition and spillover. When providers extend a product to an entire market, it is challenging to create control groups.
- **Engagement time:** Many organizations that collaborate with digital finance providers have short engagement periods and therefore fewer prospects to engage in longer-term research designs.

This complexity underlines the need to unpack the “black box” of digital finance. Digital finance impact evaluations need a lens that acknowledges complexity instead of rejecting it. A wide array of research approaches generate a useful level of evidence, but few providers use the full suite of tools available.

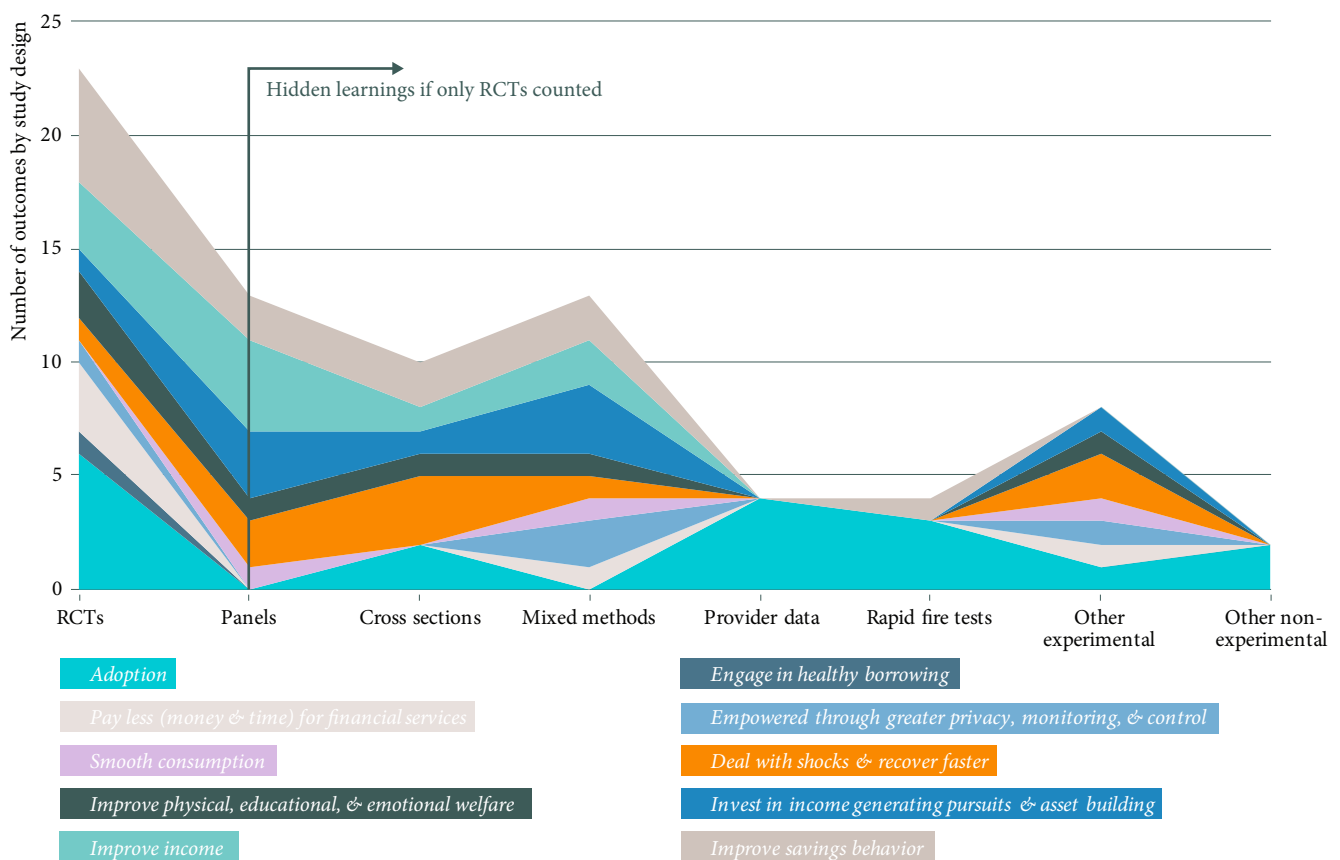
# Impact insights from diverse sources are valuable and advance the community

## When only RCTs count we lose

Experimental methods, including both random assignment and control groups, have long reigned with the “gold standard” title. But some researchers caution that when design is the primary focus, providers and researchers give inadequate attention to explaining what works, for whom, in what context, and why.<sup>3</sup> An insistence on RCTs counting as *real* evidence paralyzes efforts—go big or go home. Moreover, a synthesis of

impact evidence that included only RCTs would restrict learning opportunities. Including studies that utilize an array of methods contributes significantly to the sector’s knowledge at a point when evidence is limited. Our best sources of evidence come from numerous methodologies in dialogue with each other.

When various research designs count



Source: Analysis of the [Digital Finance Evidence Gap Map](#)

<sup>3</sup> Nutley, Powell, and Davies, “What Counts as Good Evidence? Provocation Paper for the Alliance for Useful Evidence”; Stame, “What Doesn’t Work? Three Failures, Many Answers”; Elliot Stern et al. “Broadening the Range of Designs and Methods for Impact Evaluations”; Bamberger, Rugh, and Mabry, “RealWorld Evaluation: Working Under Budget, Time, Data, and Political Constraints”; Norgbey, “Debate on the Appropriate Methods for Conducting Impact Evaluation.”

# To stretch a budget, invest in a robust theory of change, then test it

If you cannot afford experimental methods, then ensure your theory of change is robust. A theory of change describes how a program is expected to contribute to change and in which conditions it might do so; that is, “*If we do X, Y will happen because...*”. Theories of change need to be carefully crafted for each program. Abstract tables or diagrams of disjointed and ambiguous outcomes do not aid impact evaluation or program design. Some signposts for a robust theory of change are:

- **Explicit:** They clearly articulate each predicted stage of change.
  - **Context rich:** They exhibit a deep awareness of the operational context, which is fundamental to understanding impact and ergo designing impact research. Understanding context also helps a program anticipate impact heterogeneity.
  - **Plausible:** The theory that the product could lead to the suggested outcomes is conceivable and agreed upon by expert stakeholders.
  - **Testable:** The theory is specific enough to enable credible testing.
  - **Living:** The theory of change is a critical reflection tool and thus needs to be referred to and updated with insights as the program is implemented.
-



# A good theory of change allows flexibility in choosing methods

Impact research based on robust theories of change guides choices about when and how to measure outcomes. This simplifies decisions on the choice of tools in the research toolbox and helps avoid the risk of research motivated by methods rather than by theory.<sup>4</sup>

Robust theories of change coupled with comparison groups and/or causal inference methods

at the same time are the ideal. However, a robust theory of change in the absence of a comparison group is still useful. Beyond these two combinations, the remaining options are less desirable.

Elements in impact research design 2

## Presence of a rigorous theory of change

		YES	NO
Presence of a comparison group or other causal inference method	YES	Ideal	Black box
	NO	Still useful	Noise

Source: FiDA

## Look for comparison groups when you can

Impact evaluations aim to determine whether a program contributed to an observed change. There are numerous approaches to inferring causality, most of which are method-neutral. One of these is a comparison group. When considering comparison groups, the *control* group associated with RCTs likely springs to mind. In RCTs, a comparison group will have the same characteristics as the group of program participants, except the comparison group does not participate. This is to mimic the “counterfactual,” that is, what would have happened without the program.<sup>5</sup> While this approach can offer a level of certainty regarding results, the assumptions around creating counterfactual situations do not always hold in digital

finance. There are several variations of comparison groups, each with their own advocates and detractors. For example:

- In randomized encouragement design, potential users are encouraged to take part in a program while another group is not. In this setting, the difference in the level of encouragement defines the treatment and comparison group.
- Realist approaches often use intra-program comparisons, that is, comparisons across different segments involved in the same program.
- Panels and cross sections derive comparisons from the main sample by recruiting or coding for users and non-users.

<sup>4</sup> White, “Theory-Based Impact Evaluation: Principles and Practice.”

<sup>5</sup> Paul J. Gertler et al., “Impact Evaluation in Practice, Second Edition.”



- Most common is a pre-post test design that measures the outcomes of clients before and after a program. This comparison assumes that if the program had never existed, the outcome for the participants would have been the same as the pre-program state.

## Even without comparison groups you can connect the dots

When comparison groups are impractical or insufficient, other methods can shed light on how an effect was produced.<sup>6</sup> Consider the way in which justice systems aim to establish causality beyond *reasonable doubt*. Theories are developed, information is gathered, and evidence is built to create a case for the initial theory and for alternative explanations.

### BOX 1 Causal inference approaches

**Process tracing** uses clues to adjudicate between alternative explanations. Four tests are used to establish whether an explanation for the change is found to be both necessary and sufficient. Process tracing checks if results are coherent with the program theory and examines whether alternative accounts can be ruled out.

*Resource*—Punton, M. & Welle, K. (2015). “Straws-in-the-wind, Hoops and Smoking Guns: What can Process Tracing Offer to Impact Evaluation?”, Centre for Development Impact (CDI).

**Modus operandi** searches for a set of conditions that are present when the cause is effective. The consistency of the trace with the predicted pattern helps prove the program theory. Any differences from the predicted pattern might disprove the program theory or open a new line of questioning.

*Resource*—Davidson, J. (2010, May). “Outcomes, impacts & causal attribution.” Presented at Anzea regional symposium, Auckland, New Zealand.

**Qualitative comparative analysis** starts with the documentation of various configurations of conditions linked with an observed outcome. These are then subjected to a minimization procedure that establishes conditions that could account for all the observed outcomes as well as their absence.

*Resource*—Berg-Schlusser, D.; De Meur, G.; Rihoux, B.; and Ragin, C. (2008). “Qualitative Comparative Analysis (QCA) as an Approach.”

This justice-system approach has inspired evaluators, resulting in numerous strategies for understanding causality. For example, imagine you are evaluating a digital savings product that uses behavioral nudges to improve savings behavior and reduce negative coping behaviors. If the data showed that there was a high uptake of the savings product and a reduction in negative coping mechanisms, this could be because of the product or it could be because good weather resulted in a better harvest that led to fewer negative economic shocks. Can this explanation be ruled out?

Some examples of causal inference approaches are shown in Box 1. Each has slightly different organizing principles, but there are two commonalities—a robust theory of change and the use of a variety of tools to build sufficient evidence to join the dots.

**Realist analysis of testable hypothesis** elicits and explains theories on what works for whom in what circumstances and in what respects, and how. These theories are then checked against the evidence and expert opinions. This involves identifying the mechanism, context, and outcome pattern in each type of hypothesis.

*Resource*—Westhorp, G. (2014) “Realist impact evaluation: an introduction.” Methods Lab. London: ODI.

**Contribution analysis** sets out to verify the theory of change while also considering other influential factors. Causality is inferred by assessing whether the original theory was sound, activities were implemented as planned, there is evidence of outcomes, and other factors influencing the program were shown not to have contributed significantly or, if they did, the relative contribution was recognized.

*Resource*—Mayne, J. (2008). “Contribution analysis: An approach to exploring cause and effect.” The Institutional Learning and Change (ILAC) Initiative.

There are other approaches such as Most Significant Change<sup>7</sup> (MSC) that can also play a role in the impact conversation. MSC gives voice to the client on what has changed for them, why that change is important to them, and what they think caused change. While MSC does not make causal inference a primary goal, it provides some information about impact and unintended impact and can be used alongside other tools to better understand impact.

For an in-depth review of causal inference methods see Rogers, “Overview: Strategies for Causal Attribution: Methodological Briefs – Impact Evaluation No. 6.”

6 Stern, et al. “Broadening the Range of Designs and Methods for Impact Evaluations”; Stame, “What Doesn’t Work? Three Failures, Many Answers”; White and Phillips, “Addressing Attribution of Cause and Effect in Small N Impact Evaluations: Towards an Integrated Framework”; Bamberger, Rugh, and Mabry, “RealWorld Evaluation: Working Under Budget, Time, Data, and Political Constraints”; Norgbey, “Debate on the Appropriate Methods for Conducting Impact Evaluation”; Rogers, “Overview: Strategies for Causal Attribution: Methodological Briefs – Impact Evaluation No. 6.”

7 Davies and Dart, “The “Most Significant Change” (MSC) Technique: A Guide to Its Use.”

## Finding a comfortable method

Approaches that were previously used in digital finance impact research are helpful in terms of getting a sense of the possible approaches from which to draw. In the table below we highlight various methods alongside their respective best fit application, use cases in digital finance, and additional resources.

The methods described below represent a primary approach and can be coupled with other tools. For example, doing a pre-post test and using a causal inference method such as contribution analysis to further explain the quantitative results. A panel could be bolstered with in-depth qualitative interviews or with the most significant change stories to incorporate clients' inputs on their versions of events. The evaluation toolbox is full of tools in which primary methods can be enhanced by other tools to confirm, refute, explain, or enrich the findings. Here we concur that, "*the methodological 'gold standard' is appropriateness of methods selected.*"<sup>8</sup>

Perhaps the most comfortable fit for impact evaluation in digital finance comes from the realist evaluation school of thought. Realist evaluation starts and ends with a theory of change and assumes that context makes a big difference in outcomes. Emphasis is thus placed on developing and explaining "context-mechanism-outcome" (CMO) patterns, that is, "Female small business owners (C) use credit (M) to grow their business (O)." A CMO configuration may pertain to either the whole program or only certain aspects. Configuring CMOs is a basis for generating and refining the theory which becomes the final product. A realist approach is appropriate for evaluating programs that seem to work but for which the *how* and *for whom* is not yet understood; those that have so far demonstrated mixed results; and those that will be scaled up to understand how to adapt the intervention to new contexts.<sup>9</sup>

### Impact research approaches used in digital finance 1

METHOD	BEST FIT	EXAMPLE IN DIGITAL FINANCE	ADDITIONAL RESOURCES
<i>Randomized Control Trials (RCTs)</i>	<p>RCTs use two or more randomized groups: one is the treatment and the other is the control. Researchers compare these two groups to assess the effects of a program.</p> <p>RCTs work best with programs that are unlikely to change during the assessment. They are useful for generating multiple treatment groups (and thus testing product variations) and examining medium to longer-term client outcomes. They can provide a high level of certainty on effects when implemented well.</p> <p>However, in isolation RCTs are not always suited to answering questions about the mechanisms of impact.</p>	Blumenstock et al., "Promises and Pitfalls of Mobile Money in Afghanistan"; Callen et al., "What Are the Headwaters of Formal Savings?"; Schaner, "The Cost of Convenience?"; Batista and Vicente, "Introducing Mobile Money in Rural Mozambique"; Blumenstock et al., "Mobile-Izing Savings with Automatic Contributions."	Gertler et al., "Impact Evaluation in Practice, Second Edition."  Esther Duflo, "Using Randomization in Development Economics Research: A Toolkit."
<i>Randomized encouragement trials (RETs)</i>	<p>RETs are a variation of RCTs wherein participants are randomly allocated to an <i>encouragement</i> group, but can choose to participate.</p> <p>RETs have the same pros and cons as RCTs. One differentiator is that RETs are appropriate for programs for which adherence is impractical or unfair. For example, an organization might make the product accessible market-wide, but only promote it in specific localities for a period, thus creating a control and treatment group.</p>	Romero and Nagarajan, "Impact of Micro-Savings on Shock Coping Strategies in Rural Malawi."	Gertler et al., "Impact Evaluation in Practice, Second Edition."
<i>Rapid Fire Tests (RFTs)</i>	<p>In RFTs, participants are placed in random groups and exposed to product variations or messaging over a short period of time. RFTs mainly rely on administrative data such as transaction records.</p> <p>RFTs use big datasets, are relatively affordable and yield results quickly. They are excellent for testing product design tweaks. And, as they use administrative data, RFTs test first-order outcomes such as uptake and usage.</p> <p>RFTs are often context specific and thus have limited generalizability. They are not suited to questions of welfare effects due to time and data constraints.</p>	Juntos Global, "The Tigo Pesa – Juntos Partnership: Increasing Merchant Payments through Engaging SMS Conversations"; Koning, "Customer Empowerment through Gamification Case Study: Absa's Shesha Games"; Valenzuela et al., "Juntos Finanzas – A Case Study."	IPA, "Goldilocks Deep Dive: Introduction to Rapid Fire Operational Testing for Social Programs."

8 Norgbey, "Debate on the Appropriate Methods for Conducting Impact Evaluation."

9 Westhorp, "Realist Impact Evaluation: An Introduction."

<p><b>Panels</b></p>	<p>Panels are longitudinal studies in which the unit of analysis is observed at fixed intervals, usually years. Most are designed for quantitative analysis and sub-sample to create comparison groups.</p> <p>Panels can advance theory around possible effects in situations where products are market-wide and cannot be randomized or withheld. They have the potential to detect change in long-term outcomes.</p> <p>Panels often need large samples and can be labor intensive. If applied in isolation, panels are not well-positioned to explore the mechanisms of impact.</p>	<p>Suri and Jack, “The Long-Run Poverty and Gender Impacts of Mobile Money”; Jack and Suri, “Risk Sharing and Transactions Costs”; Sekabira and Qaim, “Mobile Phone Technologies, Agricultural Production Patterns, and Market Access in Uganda.”</p>	<p>Lavrakas, “Encyclopedia of Survey Research Methods.”</p>
<p><b>Cross sections</b></p>	<p>Cross section studies capture sample variables at a single point in time. Researchers can compare samples of users/non-users. They offer observations on the probable effects when time and budget are limitations.</p> <p>However, there are limitations for causal inference: While sophisticated analytical techniques can control for differences among sub-samples of product users/non-users, whether the differences occurred pre-adoption or because of adoption is uncertain. To improve causal inference, other approaches could be applied in tandem.</p>	<p>Kirui et al., “Impact of Mobile Phone-Based Money Transfer Services in Agriculture: Evidence from Kenya”; Murendo and Wollni, “Mobile Money and Household Food Security in Uganda.”</p>	<p>Lavrakas, “Encyclopedia of Survey Research Methods.”</p>
<p><b>Mixed methods</b></p>	<p>Mixed method designs collect and analyze both quantitative and qualitative data. The central premise is that a combined approach provides more robust insights than either alone.</p> <p>Mixed methods work well when initiating impact enquiries rather than resolving them. For example, in the early exploration of a new product or a new market, where impact research is limited. This more open approach allows researchers to unearth effects they had not yet considered and gives users a voice in explaining the mechanisms of the impact. When done well, mixed method designs can provide insights to confirm, refute, explain, or enrich findings.</p> <p>There are limitations to mixed methods, related to sample generation, bias, and generalizability.</p>	<p>Morawczynski, “Exploring the Usage and Impact of ‘transformational’ Mobile Financial Services”; Plyler et al., “Community-Level Economic Effects of M-PESA in Kenya.”</p>	<p>Bamberger, “Innovations in the Use of Mixed Methods in Real-World Evaluation.”</p> <p>Carvalho and White, “Combining the Quantitative and Qualitative Approaches to Poverty Measurement and Analysis: The Practice and the Potential.”</p>
<p><b>Action research</b></p>	<p>Action research studies participants’ responses to a set of mutually proposed solutions.</p> <p>Action research is suited to assessing the potential of pilot projects at the ideation stage before deciding to scale. Time frames are often short and sample sizes small. They are suited to understanding uptake and usage rather than long-term welfare outcomes.</p>	<p>Aker and Wilson, “Can Mobile Money Be Used to Promote Savings?”</p>	<p>O’Brien, “An Overview of the Methodological Approach of Action Research.”</p>
<p><b>Administrative data</b></p>	<p>Administrative data is data that tracks the progress of a product or intervention.</p> <p>Administrative data can offer insights into how clients interact with a product and may also track demographic variables that can be used to improve segmentation.</p> <p>Beyond uptake, outcomes such as savings and borrowing behaviors may also be analyzed. Intra-group comparisons may be made across client segments. The advantages of this data include time and cost savings and the possibility of accurate data and large samples.</p> <p>However, socio-economic variables of clients are rarely collected, which limits what can be asked about long-term outcomes.</p>	<p>Fenix International, “Scaling Pay-as-You-Go Solar in Uganda”; Zetterli, “Can Phones Drive Insurance Markets? Initial Results from Ghana.”</p>	<p>Caire et al., “IFC Handbook on Data Analytics and Digital Financial Services.”</p>

# Communication and certainty of impact insights

## Communicating impact insights

Qualitative or mixed method impact research is usually first on the scene and thus should be communicated early and clearly. While qualitative methods may not close the book on a topic, they are great for beginning enquiries and inviting others to follow with more intensive designs. For example, in the early days of M-Pesa, two studies using mixed methods unearthed several insights into the potential effects and use cases of M-Pesa years before experimental designs were published on the effects of M-Pesa.<sup>10</sup>

When we think about impact we often default to an assumption of *positive* impact, likely because the majority of what is published focuses on positive results. Yet, if you have tested a robust theory of change, results that do not support it are also worth sharing and discussing. While, it may be difficult, if enough of the digital finance community reported negative or null findings, the stigma of reporting such results could be mitigated, ultimately allowing us to innovate and move forward at greater speed.

Given the scarcity of evidence, the diversity of digital finance products, and the heterogeneity of the clients who use them, these signals are important and should be communicated with the proper caveats, that is, researchers must disclose the methodology and describe how/if causality was established. In the absence of *ideal* research conditions, be modest.

## Certainty of impact insights

Academia's dominance over research has made us cautious about statements of causality. While caution is good, methods—and the way we talk about the certainty of our conclusions—should meet practitioners' needs, not academic conventions.<sup>11</sup> This may not involve a numerical estimate of an effect size, but may mean interpreting it in terms of a scale of magnitude (i.e., from no obvious impact through to high impact) as well as necessity and sufficiency. That is, was the program sufficient in itself to produce the outcome, or was it a necessary but individually insufficient factor in producing the observed outcome?<sup>12</sup>

10 Morawczynski, "Exploring the Usage and Impact of 'transformational' Mobile Financial Services"; Plyler, Haas, and Nagarajan, "Community-Level Economic Effects of M-PESA in Kenya."

11 Davidson, "Causal Inference — Nuts and Bolts."

12 White and Phillips, "Addressing Attribution of Cause and Effect in Small N Impact Evaluations: Towards an Integrated Framework?"

# Go forth and join the conversation

If the digital finance community is to be guided by evidence of what works, for whom, and how, it is important to have several different tools capable of addressing questions of causal inference in a variety of contexts. Below we summarize the potential implications for future impact research design.

- **Stronger theories allow greater flexibility in methods:**

Impact research depends on rigorous theorization.

Those interested in understanding impact need to look at why we have a given product and what this product can do. Our community must also listen to clients' perspectives when trying to understand impact rather than looking for what we want to find.<sup>13</sup>

Just as digital finance products are diverse and nuanced, so too are the methods for understanding the impact of a given product for a given population in a given context. Understanding impact in the digital finance sector will never be "one size fits all." Rather, understanding impact is about finding a comfortable way to test the theory of change, given resources, program setup, and research questions. Further, the more robust theories of change will benefit from a wider variety of approaches to measuring impact.

- **Find a *comfortable* method:** Regardless of whether an evaluation has a comparison group, developing, explaining, or eliminating alternative explanations to the research findings is an important, but underused, process. As mentioned above, there are several ways to do this and it facilitates integrity and transparency and provides guidance to the reader on how they should view the results. Even if rival explanations cannot be eliminated, that information will be useful so long as the theory of change is robust. The objective is not to prove that X program is the only cause of Y effect, but that it was likely (or not) to have been part of the cause.

- **You might never be certain but you can get close:** In realist research there is no such thing as *final* truth or knowledge.<sup>14</sup> Nonetheless, it is possible to work towards a closer understanding of whether, how, and why programs work, even if we can never attain absolute certainty or provide definitive proof. Moreover, to approach the truth, we must acknowledge not only what worked but also what did not work. Insights into the positive, negative, and null effects all serve to refine theories of change, the implications of which are broader than any single program.

Impact research using experimental designs has provided great insights into the digital finance community, but that does not mean that experimental designs are the only means of revealing the effects of a program. While there may be discomfort around publishing findings that were not derived from experimental designs, we have highlighted several non-experimental impact pieces that have illuminated the effects of various products for the digital finance community. Lead with rigorous theories of change, find the comfortable research method, and join the impact conversation. Individually a single impact insight may appear insignificant, but in aggregate they advance the community of practice and our learning.

<sup>13</sup> Zollmann, "Stuff Matters: Rethinking Value in Asset Finance."

<sup>14</sup> Westhorp, "Realist Impact Evaluation: An Introduction."

- 1 *Synthesis of Evidence:*  
**Digital Finance Evidence Gap Map**
- 2 *Evidence hierarchies:* Nutley, Sandra, Alison Powell, and Huw Davies.  
“What Counts as Good Evidence? Provocation Paper for the Alliance for Useful Evidence.”  
London: Alliance for Useful Evidence, 2013.
- 3 *Theory:* White, Howard.  
“Theory-Based Impact Evaluation: Principles and Practice.” *Journal of Development Effectiveness* 1, no. 3 (September 15, 2009): 271–84.
- 4 *Mixed methods:* Carvalho, Soniya and White, Howard.  
“Combining the Quantitative and Qualitative Approaches to Poverty Measurement and Analysis.”  
WB Technical Paper 366, Washington DC: World Bank, 1997.
- 5 *Experimental methods:* Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch.  
“Impact Evaluation in Practice.”  
2nd ed. The World Bank, 2016.
- 6 *Working with constraints in evaluation:*  
Bamberger, Michael, Rugh Jim, and Mabry Linda.  
“Real World Evaluation — Working Under Budget, Time, Data, and Political Constraints, 2nd edition,”  
2011.
- 7 *Causal attribution strategies:* Rogers, Patricia.  
“Overview: Strategies for Causal Attribution: Methodological Briefs — Impact Evaluation No. 6.”  
Methodological Briefs, 2014.
- 8 *Causal attribution for mixed methods:* White, Howard, and Daniel Phillips.  
“Addressing Attribution of Cause and Effect in Small N Impact Evaluations: Towards an Integrated Framework.”  
New Delhi: International Initiative for Impact Evaluation, 2012.
- 9 *Options for causal inference:* Jane Davidson's overview of **options for causal inference** in a 20 minute webinar in the American Evaluation Association's Coffee Break series.
- 10 *Realist evaluation approach:* Westhorp, Gill.  
“**Realist Impact Evaluation: An Introduction.**”  
ODI, 2014.

# 10 Must Reads in this space



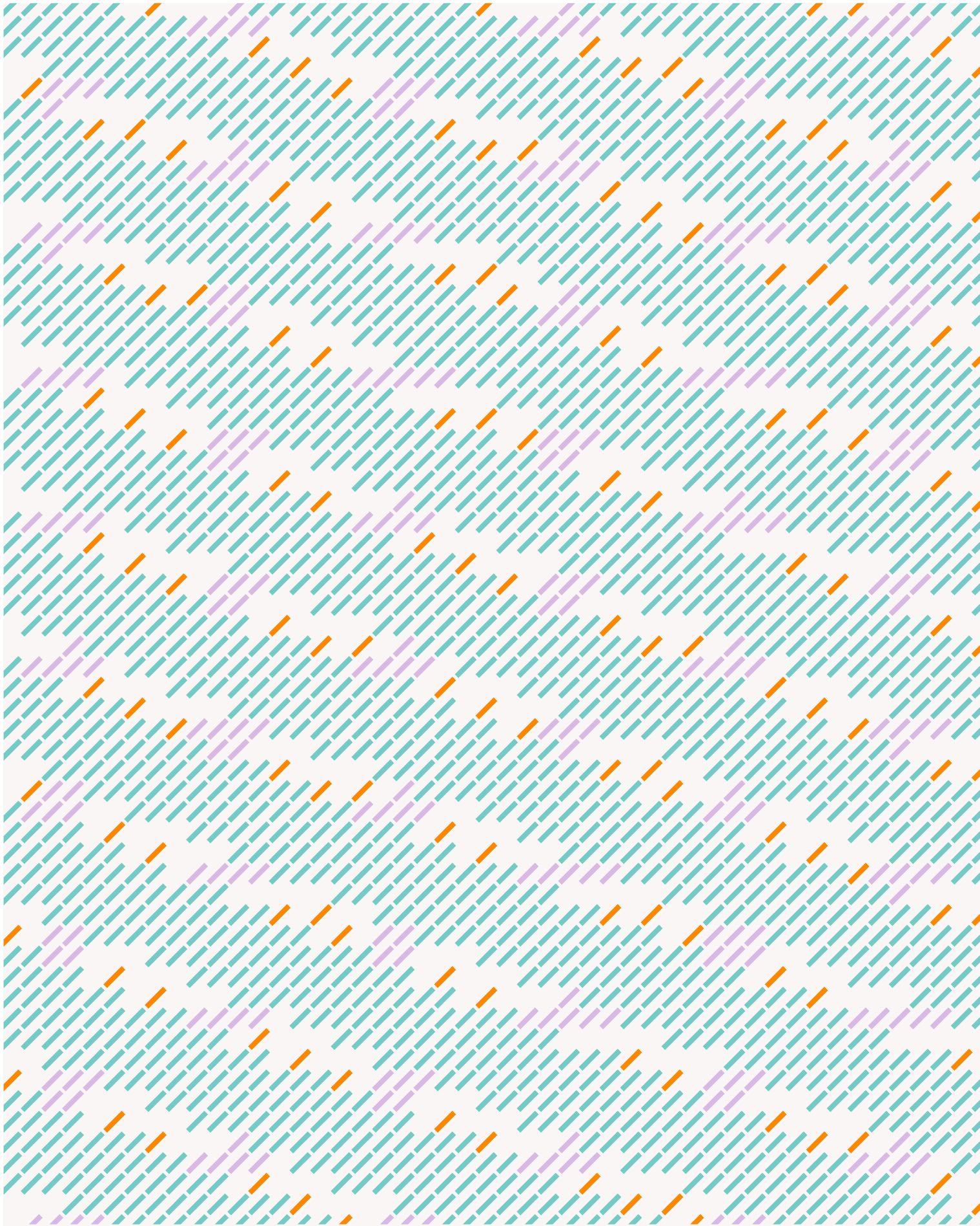
# Bibliography

- Aker, Jenny C., and Kimberley Wilson. "Can Mobile Money Be Used to Promote Savings? Evidence from Northern Ghana," 2013. [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2217554](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2217554).
- Bamberger, Michael. "Innovations in the Use of Mixed Methods in Real-World Evaluation." *Journal of Development Effectiveness* 7, no. 3 (July 3, 2015): 317–26.
- Bamberger, Michael, Jim Rugh, and Linda Mabry. "Real World Evaluation: Working Under Budget, Time, Data, and Political Constraints." SAGE Publications, 2011.
- Barry, Niamh, Donner, Jonathan and Schiff Annabel. "FiDA Snapshot 16: Digital Finance Impact Evidence Summary." Caribou Digital Publishing, November 2017. <http://www.financedigitalafrica.org/snapshots/16/2017/>.
- Batista, Cátia, and Pedro C. Vicente. "Introducing Mobile Money in Rural Mozambique: Evidence from a Field Experiment." *NovaAfrica*, June 2013. [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2384561](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2384561).
- Blumenstock, Joshua E., Michael Callen, Tarek Ghani, and Lucas Koepke. "Promises and Pitfalls of Mobile Money in Afghanistan: Evidence from a Randomized Control Trial." In *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*, 15. ACM, 2015.
- Blumenstock, Joshua E., Michael Callen, and Tarek Ghani. "Mobile-Izing Savings with Automatic Contributions: Experimental Evidence on Present Bias and Default Effects in Afghanistan." CEPR Discussion Paper No. DP11400, July, 2016. <https://papers.ssrn.com/abstract=2814075>.
- Caire, Dean, Leonardo Camiciotti, Soren Heitmann, Susie Lonie, Christian Racca, Minakshi Ramji, and Qiuyan Xu. "IFC Handbook on Data Analytics and Digital Financial Services." IFC, 2017. <https://www.ifc.org/wps/wcm/connect/22ca3a7a-4ee6-444a-858e-374d88354d97/IFC+Data+Analytics+and+Digital+Financial+Services+Handbook.pdf?MOD=AJPERES>.
- Callen, Michael, Suresh De Mel, Craig McIntosh, and Christopher Woodruff. "What Are the Headwaters of Formal Savings? Experimental Evidence from Sri Lanka." Cambridge, MA: National Bureau of Economic Research, 12/2014. <http://www.nber.org/papers/w20736.pdf>.
- Carvalho, Soniya, and Howard White. "Combining the Quantitative and Qualitative Approaches to Poverty Measurement and Analysis: The Practice and the Potential." World Bank Publications, 1997.
- Conrad, Murendo, and Meike Wollni. "Mobile Money and Household Food Security in Uganda." Department of Agricultural Economics and Rural Development, Georg-August-University of Goettingen, 37073 Goettingen, Germany, 2016. [http://ageconsearch.tind.io/record/229805/files/GlobalFood\\_DP76.pdf](http://ageconsearch.tind.io/record/229805/files/GlobalFood_DP76.pdf).
- Davidson, Jane. "Causal Inference – Nuts and Bolts." May 21, 2009. <http://realevaluation.com/pres/causation-anzea09.pdf>.
- Davies, Rick, and Jess Dart. "The 'Most Significant Change' (MSC) Technique: A Guide to Its Use." Rick Davies, 2007.
- Duflo, Esther, Rachel Glennerster, and Michael Kremer. "Using Randomization in Development Economics Research: A Toolkit." *Centre for Economic Policy Research*, 2007. <https://economics.mit.edu/files/806>.
- Fenix International, "Scaling Pay-as-You-Go Solar in Uganda." Mobile for Development, December 14, 2015. <https://www.gsma.com/mobilefordevelopment/programme/m4dutilities/fenix-international-scaling-pay-go-solar-uganda>.



- Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch. *Impact Evaluation in Practice*, 2nd. ed. The World Bank, 2016.
- IPA. “Goldilocks Deep Dive: Introduction to Rapid Fire Operational Testing for Social Programs.” IPA, 2016. [https://www.poverty-action.org/sites/default/files/publications/Goldilocks-Deep-Dive-Introduction-Rapid-Fire-Operational-Testing-for-Social-Programs\\_1.pdf](https://www.poverty-action.org/sites/default/files/publications/Goldilocks-Deep-Dive-Introduction-Rapid-Fire-Operational-Testing-for-Social-Programs_1.pdf).
- Jack, William, and Tavneet Suri. “Risk Sharing and Transactions Costs: Evidence from Kenya’s Mobile Money Revolution.” *The American Economic Review* 104, no. 1 (January 1, 2014): 183–223.
- Juntos Global. “The Tigo Pesa — Juntos Partnership: Increasing Merchant Payments through Engaging SMS Conversations.” Juntos Global, 2015. <http://juntosglobal.com/wp-content/uploads/2015/12/Tigo-Pesa-Juntos-Merchant-Payments-Partnership.pdf>.
- Kirui, Oliver, Julius Okello, Rose Nyikal, and Georgina Njiraini. “Impact of Mobile Phone-Based Money Transfer Services in Agriculture: Evidence from Kenya,” 2013. [http://ageconsearch.umn.edu/bitstream/173644/2/3\\_Kirui.pdf](http://ageconsearch.umn.edu/bitstream/173644/2/3_Kirui.pdf).
- Koning, Antonique. “Customer Empowerment through Gamification Case Study: Absa’s Shesha Games.” CGAP, 2015. <http://www.cgap.org/sites/default/files/Working-Paper-Customer-Empowerment-Through-Gamification-Absa-Shesha-Dec-2015.pdf>.
- Lavrakas, P. J. *Encyclopedia of Survey Research Methods*. SAGE Publications Ltd, 2008.
- Morawczynski, Olga. “Exploring the Usage and Impact of ‘transformational’ Mobile Financial Services: The Case of M-PESA in Kenya.” *Journal of Eastern African Studies* 3, no. 3 (2009): 509–25.
- Norgbey, Enyonam B. “Debate on the Appropriate Methods for Conducting Impact Evaluation.” *Journal of Multidisciplinary Evaluation* 12, no. 27 (2016). [http://journals.sfu.ca/jmde/index.php/jmde\\_1/article/download/454/422](http://journals.sfu.ca/jmde/index.php/jmde_1/article/download/454/422).
- Nutley, Sandra, Alison Powell, and Huw Davies. “What Counts as Good Evidence? Provocation Paper for the Alliance for Useful Evidence.” London: Alliance for Useful Evidence, 2013.
- O’Brien, Rory. “An Overview of the Methodological Approach of Action Research,” 1998. <http://web.net/~robrien/papers/xx%20ar%20final.htm>.
- Plyler, Megan G., Sherri Haas, and Geetha Nagarajan. “Community-Level Economic Effects of M-PESA in Kenya.” Iris Center, University of Maryland, 2010. <http://www.fsassessment.umd.edu/publications/pdfs/Community-Effects-MPESA-Kenya.pdf>.
- Rogers, Patricia. “Overview: Strategies for Causal Attribution: Methodological Briefs — Impact Evaluation No. 6.” Methodological Briefs, 2014. <https://ideas.repec.org/p/ucf/metbri/innpub751.html>.
- Romero, Jose, and Geetha Nagarajan. “Impact of Micro-Savings on Shock Coping Strategies in Rural Malawi.” IRIS Centre, 2011. <http://www.fsassessment.umd.edu/publications/pdfs/iris-shocks-bill-melinda-gates-foundation.pdf>.
- Schaner, Simone. “The Cost of Convenience? Transaction Costs, Bargaining Power, and Savings Account Use in Kenya,” 2016. <http://www.poverty-action.org/sites/default/files/publications/The-Cost-of-Convenience-April-2016.pdf>.
- Sekabira, Haruna, and Martin Qaim. “Mobile Phone Technologies, Agricultural Production Patterns, and Market Access in Uganda,” 09/2016. <http://ageconsearch.umn.edu/bitstream/246310/2/66.%20Mobile%20phone%20technologies.pdf>.
- Stame, Nicoletta. “What Doesn’t Work? Three Failures, Many Answers.” *Evaluation* 16, no. 4 (October 1, 2010): 371–87.
- Stern, Elliot, Nicoletta Stame, John Mayne, Kim Forss, Rick Davies, and Barbara Befani. “Broadening the Range of Designs and Methods for Impact Evaluations.” DFID, 2012. <https://www.oecd.org/derec/50399683.pdf>.

- Suri, Tavneet, and William Jack. "The Long-Run Poverty and Gender Impacts of Mobile Money." *Science* 354, no. 6317 (December 9, 2016): 1288–92.
- Valenzuela, Myra, Nina Holle, and Wameek Noor. "Juntos Finanzas – A Case Study." CGAP, October 2015. <http://www.cgap.org/sites/default/files/Working-Paper-Juntos-Finanzas-A-Case-Study-Oct-2015.pdf>.
- Westhorp, Gill. "Realist Impact Evaluation: An Introduction." ODI, 2014. <https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9138.pdf>.
- White, Howard. "Theory-Based Impact Evaluation: Principles and Practice." *Journal of Development Effectiveness* 1, no. 3 (September 15, 2009): 271–84.
- White, Howard, and Daniel Phillips. "Addressing Attribution of Cause and Effect in Small N Impact Evaluations: Towards an Integrated Framework." New Delhi: International Initiative for Impact Evaluation, 2012.
- Zetterli, Peter. "Can Phones Drive Insurance Markets? Initial Results from Ghana." CGAP, February 2013. <http://www.cgap.org/blog/can-phones-drive-insurance-markets-initial-results-ghana>.
- Zollmann, Julie. "Stuff Matters: Rethinking Value in Asset Finance." CGAP, September 5, 2017. <http://www.cgap.org/blog/stuff-matters-rethinking-value-asset-finance>.



*Partnership for* **FINANCE**  
*in a* **DIGITAL AFRICA**