World Bank Africa Gender Innovation Lab (GIL)  
Mid-term Review

Mayra Buvinic and Rebecca Furst-Nichols¹  
(November 2015)

Abstract

This mid-term review of the World Bank Africa Gender Innovation Lab (GIL) used relevant gender and development and impact evaluation literature, results of two GIL-administered short surveys, and 30 individual interviews plus a team group interview to examine GIL’s main outputs and assess how well GIL’s processes and activities are structured to facilitate research uptake and influence policy regarding gender issues in Africa.

Main findings include: (1) The quality of GIL’s evaluation research is very high, and generating rigorous evidence on what works to increase women’s productivity and incomes is a major value-added of GIL’s work. (2) Fostering commitment to evidence-based policymaking is done through in-country evaluation workshops and the fellows program. The workshops are rated highly but can be fined tuned, and the fellows program needs better organization and expanded recruitment. (3) There is clear evidence that GIL’s research has been used to modify policies, scale-up (or down) innovations, and influence policies indirectly. A conservative staff calculation based on 29 projects is that for every GIL dollar spent, 46 project dollars are being directly influenced. (4) The ‘Leveling the Field’ synthesis report combined all the features of good evidence for policymaking, which helps explain its success in research uptake. Further promotion of research uptake awaits a critical mass of results which is not yet available.

GIL has a flat reporting structure that has evolved successfully but may need a mid-level management position as the team grows in size. The budget is on track and managed competently, and GIL is on track to meet the goals set in the proposal to DFID. The text of this review contains a number of specific suggestions going forward. More generally, it is clear that the Africa GIL is a very high quality, well-run evaluation research initiative and is poised for continued growth. This growth, however, should be managed carefully and build on GIL’s uniqueness in researching what works to promote women’s economic empowerment, a ‘public good’ type activity that is otherwise underfunded in the Bank and externally.

¹ Mayra Buvinic (mayra.buvinic@gmail.com) and Rebecca Furst-Nichols (rfurst.nichols@gmail.com) prepared this report as independent consultants to the World Bank.
# Table of Contents

*Framework* ................................................................. 2

*GIL’s History* .............................................................. 3

*Quality and Value-Added of GIL’s Research* ....................... 6

*Fostering Commitment to Evidence-based Policymaking* .......... 10

*GIL’s Effects on Policy* .................................................. 12

*Promoting Research Uptake* ........................................... 14

*GIL’s Organizational Structure, Management and Budget* ......... 15

*Conclusions and Suggestions Looking Forward* .................... 18

*Table 1: Sample of GIL’s Impact Evaluations*

*Table 2: Progress Toward Meeting Quantitative Targets*

*Table 3: List of People Interviewed*
Framework

The World Bank’s Africa Gender Innovation Lab (GIL) does evaluation research in Sub-Saharan Africa to build rigorous evidence on what works to empower women economically and address gender inequality in the productive sectors.

This mid-term review examines how well GIL’s processes and activities are structured to facilitate research uptake and influence policy regarding gender issues in the region.

Research has conceptual, capacity building, and instrumental effects on policy. That is, research influences policy by: (a) broadening policy horizons, (b) expanding policy capacities, and (c) affecting policy regimes.²

Following these three different types of policy influences, the review examines, first, the quality and value added of the concepts and ideas generated by GIL to shift the policy dialogue and foster strengthened knowledge of effective gender policies; second, the commitment to evaluation and to evidence-based policymaking on gender that GIL generates by nurturing new talent and expanding capacities; and third, GIL’s capacity to affect policies and programs directly, or research’s instrumental value.

To do this, the review examines the three major GIL outputs: impact evaluation research, impact evaluation workshops and fellowships, and retail advice and support to projects (through the impact evaluation work embedded in projects, project technical assistance, and integration into project teams). It also examines GIL’s communication and publication activities to assess the process by which policy makers, practitioners and other development actors become aware of and access GIL’s research outputs (research uptake).³

Factors that mediate the effective intersection of the supply of knowledge and the demand for it (uptake) by policymakers include the effectiveness of researchers in communicating this knowledge to policymakers; an enabling environment for evidence-based policy making, including the capacity of policy makers to use the knowledge and policy makers’ own beliefs; and the nature of the policy change that is supported by the evaluation research. Incremental policy changes are easier to adopt than transformative ones, since they are less likely to challenge dominant interests.⁴ Assessment of research uptake needs to allow for a sufficiently long time frame with which to understand the influence of research.

GIL’s History

GIL grew organically in the World Bank Africa region to counter rhetoric on gender issues with rigorous evidence on what works.

---

The GIL was conceived as a response to a frustration in the Africa region with much rhetoric on gender and development, and attempts at ‘gender policing,’ but little robust evidence on what worked or practical advice to Bank teams on what to do to integrate gender issues into operations. The PREM (Poverty Reduction and Economic Management) Africa Director at the time, circa 2007, brought in a capable researcher working in the PREM anchor for poverty to share his time between PREM Africa and the research department (DEC), and assigned him to start a small impact evaluation gender initiative linked to operations. The guiding principles were rigor and a focus on interventions.

Donor support fostered GIL’s growth and the publication of the 2012 World Development Report provided the conceptual underpinnings.

A Bank-wide initiative, the Gender Action Plan (GAP), which had been launched in 2006 with a similar purpose and a focus on the economic sectors, provided the initial funding through an umbrella, multi-donor trust fund that grew in size quickly and financed many of the early GIL impact evaluations. Additional funding sources in this first period were the Bank Netherlands Partnership (BNPP) trust fund and TFESSD (TF for environmentally and socially sustainable development). The 2012 World Development Report on Gender Equality and Development gave the GIL added momentum. The Report presented strong business and fairness arguments for addressing gender inequalities and stressed the multiplicative effects of empowering women economically. It examined the determinants of gender inequalities, providing a fuller intellectual architecture for GIL’s work moving forward. However, it also revealed how little rigorous knowledge there was about what worked, especially in the economic sectors, in agriculture, and in Africa. Two other publications around the same time underscored these same gaps.

GIL’s second phase (2.0) took-off in 2013, thanks to generous support from DFID channeled through the Umbrella Facility for Gender Equality (UFGE). In preparation for this grant, GIL developed clear objectives, a framework and indicators, and a theory of change. GIL’s focus on evaluation research seeking ‘proof of concept’ (internal validity) or ‘generalizations’ (external validity) of innovative interventions in the productive economic sectors is strategic in addressing gender knowledge gaps and poised to provide valuable, practical guidance for Bank operations.

GIL 2.0 builds on and brings forward more systematically the work started during GIL 1.0 and shows the iterative way that useful policy evidence builds around a critical mass of impact evaluations.

GIL 2.0 builds on and brings forward more systematically the work started during GIL 1.0. Highlights of the first phase include the Liberia adolescent girls’ initiative project, EPAG. EPAG tested a demand-driven vocational training program that had proven successful with young women in higher-income Latin America in a low income, post-conflict African country. With some variations, including focusing training more on entrepreneurship than on wage work, results were convincing (young women increased their earnings significantly) showing that the model was generalizable in a very different economic environment.

5 As of last fiscal year, the GIL leader is a full time staff member in AFR.
EPAG, in turn, triggered additional evaluation work testing ‘second order’ questions on youth vocational programs, such as the value of providing information to young women on earnings from different trades (in youth skills training projects in the Republic of Congo and Benin) and of life skills mentoring with adolescent girls (IRC’s Sisters of Success Project in Liberia). Answers to these questions could further fine-tune and strengthen the impact of the youth training model first piloted in Latin America, and show the iterative, interconnected way that useful policy evidence builds around a critical mass of project impact evaluations. About half a dozen rigorous evaluations of ‘jóvenes’ programs in Latin America preceded EPAG in Liberia.

EPAG influenced the design of the Bank’s Republic of Congo Skills Development for Employability Project, with its focus on the employment needs of young women and stronger emphasis on business rather than jobs skills, as well as additional financial support for young mothers in the program. It is worth noting that EPAG has had the endorsement from Liberian President Johnson-Sirleaf from the outset, underscoring the important role of high-level political support for gender initiatives.

**GIL-Research Group findings substituted for gender-informed diagnostics and were used to avoid a negative impact on women of a well-intended but gender-blind program design.**

Another highlight of GIL 1.0 is the DFID funded Rwanda land title registration project. A GIL and Research Group team evaluation of a Land Tenure Regularization pilot that mandated joint (husband and wife) urban and rural land titles found significant increases in land investment, larger for women (19 percent versus 10 percent for men), when they had access to a title. The evaluation, however, also found that unmarried women living in common-law unions (around 30 percent of the households in the pilot) lost out since the title program had not contemplated common-law wives. Thanks to the evaluation results, the national land-titling program, which issued 6.1 million titles, was redesigned, granting co-owner titles to both married and unmarried women. This is the first policy impact of GIL’s evaluation work (recorded in 2013).

The GIL-Research Group team’s findings avoided a negative impact on women from what would have been an otherwise well-intended but poorly designed land registration program from a gender perspective. The evaluation substituted for what should have been good, gender-informed diagnostics in the Rwanda land titling registration program in the first place. While this was not a Bank-funded project, it nevertheless raises the issue of the role and scope of GIL evaluations versus other Bank diagnostic instruments for gender-informing the lending portfolio.

**GIL now sits at the intersection of two Bank work streams, Africa and gender, adding administrative costs but increasing its potential range of influence.**

With the Bank’s reorganization in 2013-2014, the Africa region Vice President, convinced of GIL’s value to inform operations, played an important role in keeping GIL in the region and continues to champion and support it. GIL now sits in Africa’s chief economist office; some staff report to and are covered by the Gender Cross Cutting Solution Area (CCSA) but work for GIL full time. This hybrid model has required sorting out administrative hurdles and reporting lines, but on the other hand has enabled interaction with two different work streams at the Bank (Africa and gender), increasing GIL’s visibility and exposure to internal clients, and its access to core services such as communications support. This hybrid model has also helped with external clients, facilitating interactions at the country-level and with global gender actors. The Gender CCSA’s understanding of the importance of gender impact evaluation work (facilitating fast reviews and approvals) combined with Africa’s regional expertise has worked well for the GIL team.
GIL has influenced the organization of Bank-wide gender work, as evidenced by the gender strategy under preparation.

Following the Africa GIL’s model, the Bank’s gender strategy, currently under development, proposes a federal system of gender labs covering the different regions, with a coordinating council that is able to transfer knowledge and synthesize the body of gender evidence across Bank regions. The labs would both continue to investigate what works and also do deep-dive country diagnostic work up-front, therefore having a more rounded, overall influence on the program cycle.

In addition to supporting gender innovations and conducting impact evaluations, the Africa GIL has strived to communicate research results, build capacity on rigorous impact evaluation work, and influence policy. A staff member on a development assignment from what is now the Gender CCSA developed a comprehensive research uptake strategy for GIL’s work.

In the sections that follow, we examine the different GIL activities and outputs following the four dimensions outlined in the framework for this mid-term review:

- Quality and value-added of the concepts generated
- Fostering commitment to evidence-based policymaking
- Research’s direct effects on policy (instrumental value)
- Promoting research uptake

In addition, we review GIL’s organizational structure, management, and budget. We end the review with summary conclusions and suggestions for GIL looking forward. The list of people interviewed for the review is included in Table 3. The results of two GIL administered surveys, a mid-term short survey of clients (with 32 respondents) and an internal 360 management survey, informed our analysis, as well as the relevant gender and development and impact evaluation literature.7

Quality and Value-Added of GIL’s Research

Quality
The quality of GIL’s evaluation research is very high, based on reviewing a random sample of GIL 1.0 evaluation designs and GIL 2.0 concept notes.

We assessed the quality of GIL’s ongoing evaluation research by reviewing a random sample of the GIL portfolio of 53 evaluations and checking to see if the concept notes satisfactorily addressed the following key questions/issues in impact evaluation research: clear policy objectives; relevant constraints addressed; ability to spell out channels or mechanisms leading to impact; accounting for confounding factors and spillover effects; and addressing selection bias and heterogeneity.8

7 The electronic survey was sent to around 590 names in GIL’s mailing list. Of the 32 survey respondents, half were government representatives and half were from international agencies. 59 percent lived in Africa.

We randomly chose 4 projects from GIL 1.0 and 8 projects from GIL 2.0, and rated them on the above dimensions. See Table 1 for the list of projects and their ratings. We also reviewed the quality control process set up by GIL.

As Table 1 indicates, the quality of GIL’s ongoing research evaluations is very high. We found no fault with any of the concept papers we reviewed. The concept notes for GIL 2.0 evaluations are technically sound, thorough and rich in detail. They reflect significant, above the norm, up-front investments. GIL 1.0 evaluations did not produce GIL concept notes (this does not mean that they did not go through a thorough review process, but one that was not systematized at the GIL level), so staff recollections guided this process.

Supporting the quality of its work, GIL clients, both those doing Bank operations and external implementing partners, were uniformly positive about the GIL team’s technical competence.

A rigorous review process, building on long-standing Bank quality assurance mechanisms, helps to account for the high quality of GIL’s research.

The rigorous review process contributed to the quality of the concept notes reviewed. Concept notes receive general guidance as well as specific input from the Steering Committee, detailed feedback from 2 to 3 external technical reviewers (with written responses to the feedback), and an additional review step and sign-off from the Lead Impact Evaluation economist for Africa. It is a systematic review process that builds on the Bank’s long-standing quality assurance processes for research and operations.

**GIL may want to deepen exploration of heterogeneity in evaluation research and build common outcome metrics.**

Two dimensions related to the quality (and value added) of research outputs that GIL may want to deepen is, first, a more systematic exploration of heterogeneity among women, and especially if and how interventions work differently for different economic groups. Second, as a critical mass of impact evaluations is building, it would be good to identify a few intermediate and final outcome indicators measuring women’s economic empowerment and gender equality that can be shared and systematically tracked across projects, to help convey an overall story of GIL’s impact. This exercise can help ensure that all the different evaluations in GIL’s portfolio share common operational definitions of economic empowerment and gender equality.

**GIL should plan to follow up the evaluations in the portfolio and repeat outcome measures over time to account for transitory and delayed effects. It should also examine systematically failed interventions.**

GIL needs to plan for follow-ups to the evaluations in the portfolio to measure the sustainability of the changes observed. This is particularly important, since changes in gender outcomes as result of interventions are sometimes transitory and others emerge with a delay (gender norms are often hard and slow to change), while GIL relies heavily on RCTs which tend to favor projects that yield quick results. Another criticism of RCTs is that they reinforce a market for knowledge where only what works well gets published; failures seldom see daylight. GIL staff and colleagues in the Bank’s research group have published a paper addressing the issue of failure and, once a critical mass of results comes in, GIL should systematically explore the nature of

---

9 Ravallion, “Evaluation in the Practice of Development.”
failed interventions in its portfolio.\textsuperscript{10} Identifying and explaining failures is complex and time consuming, however, and GIL should create some space and allocate sufficient staff resources to this task.

\textit{There remains an important role for qualitative research and the existing capacity in this area should be expanded.}

While initially it was not clear how the qualitative component of the GIL would fit within the structure of the team, this role has proven to be valuable, with economist team members seeking assistance on issues such as understanding channels and mechanisms of impact. The GIL researcher in charge has training in both qualitative and quantitative methods, which has facilitated integrating qualitative methods into GIL’s evaluation studies.

Qualitative research has been used to clarify confusing evaluation findings, inform the design of survey instruments, deepen understanding of specific topics, elicit the intervening variables that may explain the changes observed in evaluation studies (such as, ‘what is it about the mentoring that lead to changes in the behavior of young women in Liberia’s Sisters of Success project?’) and, in two cases, as stand-alone studies informing the design of programs (such as studying household allocation mechanisms in polygamous households in Burkina Faso before designing a cash transfer program).

Despite the significant value of the qualitative analysis, the team has only one Qualitative Research Specialist who currently manages 12 projects. This capacity should be expanded in order to maintain and control quality. This is particularly the case since, ideally, qualitative methods should be integrated throughout the life of the evaluation project, rather than only at certain points. One idea for supporting this work is to recruit one or more junior fellows with qualitative expertise.

GIL may wish to review and reflect on how best to package and communicate the results of the qualitative research internally and for external audiences. Quantifying qualitative results makes little sense conceptually but often resonates the most with internal Bank audiences. GIL may also wish to invest in longitudinal ethnographic work to explain changes over time regarding selected gender issues with potential critical importance for the design of Bank operations, such as gender and household dynamics intervening in the distribution of government transfers.

\textbf{Value – Added}

\textit{Generating rigorous evidence on what works to increase women’s productivity and incomes is a major value-added of GIL’s work.}

GIL has four thematic areas: land, agriculture, entrepreneurship, and youth employment. Overall, the value-added of GIL is in generating rigorous evidence on what works across these areas and particularly in agriculture, the thematic area with the weakest evidence-base on what works (in contrast to the rich documentation that exists on women’s roles in agriculture and the determinants of their agricultural productivity).

The ‘Leveling the Field’ synthesis research report added valuable new knowledge by documenting empirically that providing women with access to agricultural services was not

The value-added of testing agricultural services innovations is higher than the value-added of measuring differential impacts by gender.

Based on this message, 7 of the 13 agriculture projects in the GIL portfolio modify the provision of agricultural services (seeds, extension, training, contracts to farmers, capital) to address specific constraints and increase uptake by female farmers. They include quotas for females in farmer groups, issuing farming contracts to women in the farming household, training farming couples, and incentives to female farmers to grow cash crops, among others. The value-added of these evaluations is evident and seems higher than that of others in the portfolio that instead measure the differential impact of agricultural services on female and male farmers or, more traditionally, look at nutritional and food security outcomes.

GIL includes 5 land and asset projects. Four explore the impact of providing land titles to women and how best to do it, and this includes the national rollout and scale-up of the Rwanda land title registration project. They should add value to a growing body of evidence that suggests but has yet to rigorously demonstrate that women’s land ownership increases their productivity in agriculture and ‘empowers’ them in the household and community. Future work should begin to explore the impacts of increasing women’s access to land and housing titles in urban areas – an issue largely unexplored to date.

The large numbers of ‘idle’ young women in African countries underscores the value-added of youth employment interventions targeted to young women.

The 11 youth employment projects test different interventions to increase young women’s employability. Their value-added is underscored by the large percentage of young women in Sub-Saharan Africa who are ‘idle,’ that is, are neither in school nor in the labor force, pointing out the need for effective, scalable interventions that can improve young women’s successful transition from school to employment. Particularly interesting is a project in Sierra Leone implemented by BRAC that seeks to disentangle the effects of different interventions (girls clubs, life skills training, livelihood training, credit) to identify the most binding constraints and pare down program components from a bundle of multiple interventions (the current ‘best practice’) to a few that can be more easily scaled-up.

Testing interventions promoting female entrepreneurship is more novel than documenting gender differentials.

The 15 entrepreneurship (private sector development) projects cover a range of diverse interventions, from training on taxes and information on gender-based violence, to new banking products, to business plan competitions and the benefits of formalizing SMEs. As in the case for agriculture, the value-added of projects that test interventions directed specifically to address the constraints faced by women entrepreneurs seems higher than those that document gender differentials in response to a variety of gender-neutral interventions. Especially novel are evaluating the impact of childcare services on agricultural outcomes (testing for the first time in a rural environment an intervention proven to increase women’s employment in urban areas), psychometric testing to replace credit collateral, and work on ‘cross-overs,’ or transitioning entrepreneurs from low to high return, ‘nontraditional’ women’s business activities (therefore, seeking transformative rather than incremental impacts).
GIL exposes the World Bank to novel gender interventions implemented by external partners.

‘Innovation,’ one of GIL’s main premises, is widely overused in development work. However, in the context of Bank operations, ‘innovation’ can also mean introducing new interventions into the Bank’s project portfolio. Most of the interventions mentioned above, designed specifically to overcome constraints to women’s economic participation, are all novel to Bank operations. GIL’s outreach to and work with external implementing partners has facilitated testing these innovations and increased GIL’s legitimate claim to the term. Smaller, more flexible agencies are usually better suited to implement interventions targeting women and girls, and GIL had the foresight to include them in their open call for proposals. Thirty percent of GIL’s project portfolio of 53 impact evaluations is devoted to collaborations with external partners.

GIL’s thematic synthesis reports (white papers), under preparation, can help take stock of the evidence and guide future strategic research investments.

Over time, GIL has moved from an opportunistic approach (GIL 1.0) to more systematically choosing projects and research questions to address (GIL 2.0), therefore increasing the likelihood that GIL evaluation research will advance strategic, useful gender policy knowledge. The synthesis reports (white papers) under preparation for GIL’s thematic areas are important vehicles to systematize the evidence and guide future research investments.

Fostering Commitment to Evidence-based Policymaking

The GIL intends to build capacity around evidence-based policymaking through in-country impact evaluation workshops and a Fellows program. To assess each of these intended outcomes, we interviewed client country and non-governmental implementing partners, and Fellows.

Different capacities around impact evaluation, and different goals for participating in the evaluation, should be carefully considered when designing the workshops.

All implementing partners reported having positive experiences working with the GIL team; GIL members were supportive, professional, and systematic. Impact evaluation workshops were rated highly in GIL’s client survey – 69 percent said the workshops were useful. Some government partners felt that the impact evaluation workshop material was difficult to understand and likely surpassed the ability of participants to fully grasp it. One partner suggested conducting additional workshops after the project is underway and the implementing team has a better sense of the work; at that point, they may be in a better position to understand the details. It may be useful for the GIL team to work with a professional trainer to translate the concepts for in-country workshops in a way that is understandable to implementers who are entirely new to impact evaluation concepts.

For those implementers with prior evaluation experience, usually those with organizational capacity around impact evaluation, the GIL workshops were very clear. IRC, for example, has a strong research and evaluation team at headquarters, and therefore had significant capacity to participate in the workshop and support getting the evaluation going in the field.

Some of the evaluations were mandated by donors, and implementing staff questioned the value of the results for their own purposes. This could be partly due to the differing information needs for programs as compared to research results. Impact evaluations often do not have the level of detail about the program to provide practical information for practitioners, and robust project
monitoring and evaluation (M&E) done or contracted out by the implementing agency (rather than by the GIL team) should accompany impact evaluations.

*Given the capacity-building aims of the GIL, it is of utmost importance to expand recruitment and structure around the Fellows program, and to get it right.*

The Junior and Senior Fellows program is somewhat behind schedule (with only 2 of 9 target Fellows) but plans are underway to revise the program and ramp it up. It has been more difficult than expected to recruit Senior Fellows, since they usually have a wider existing network and less available time, and therefore do not stand to benefit as much from the program. The GIL team has determined that the program will achieve its aims better through a sole focus on Junior Fellows who aim to obtain PhDs in a relevant discipline and can benefit from GIL’s training, network, and mentorship. They recognize that much can be done to expand recruitment for and organization of the Fellows program. At the same time, the program requires significant staff time and dedication, and it will be important to assess with potential mentors (most likely GIL thematic leaders) in advance of any expansion plans their willingness and ability to devote time to guiding Fellows’ professional development. Strong up-front investment on the part of the Fellow and the mentor to clarify expectations and goals will contribute to the program’s success.

The current and previous Junior Fellow each gave promising reviews of the program, with helpful suggestions for consideration as the program evolves (see below). They reported feeling logistically and academically supported, mentored, intellectually stimulated, exposed to a broad range of experiences, and integrated socially within the GIL team. One Fellow reported utilizing the relationship with his mentor, broad exposure within the Bank, and the network to advance significantly in his field. However, another suffered from significant delays in his start date, despite having already left his previous post, due to a Bank-wide reorganization occurring during his tenure.

With such a small sample, it is not possible to generalize from these experiences, but there are lessons to be drawn. As their primary recommendation, both Fellows strongly emphasized the need for clear Terms of Reference prior to beginning the Fellowship, in order to guide the work, expected contributions, and place on the team. One recommended that especially when working with young people from Africa who are not always accustomed to a culture of speaking up and advocating for themselves within institutions, a clear TOR can help build the Fellow’s confidence and ensure smooth integration into the team. Additional recommendations included:

- Determine a set timeline for the Fellowship and remain on schedule, since Fellows are balancing other professional opportunities as well as education timelines.
- While it is important for Fellows to come in with a specific area of focus and experience, it would be valuable for the GIL to facilitate the Fellow’s participation on and exposure to projects focused on other GIL thematic areas, and other World Bank teams (e.g. LSMS, DIME).
- The relationship with the mentor is central to the experience of the Fellow. The mentor should know what the Fellow intends to get out of the program and help facilitate this, and make him/herself available to get to know the Fellow personally and professionally.
- Personal initiative of the Fellow matters. It is important for the Fellow to be very specific about what he/she wants to get out of the experience, and to verbalize that early. A phone call between the mentor and the Fellow before the program begins may be useful. In the words of one interviewee, Fellows need to “help them help you”, and not be passive their needs.
• It may be helpful for prior years’ Fellows to speak by phone with incoming Fellows before their arrival to help orient them and provide ideas for getting the most out of the experience.
• It is helpful for the GIL team to introduce Fellows to opportunities available in the Washington area, such as relevant courses, and to facilitate networking and connections with universities and professors to aid in the process of applying for graduate school.
• The recruitment process should be improved upon and expanded. The current Fellows were identified in a somewhat ad hoc manner because they knew individuals on the GIL team. It would be good to expand the pool of applicants deliberately, perhaps working through local universities and former Fellows. It is important to note that both Fellows to date have been male, and a concerted effort to recruit women should be made in subsequent rounds, though we do not recommend instituting a formal quota.

With improvement and expansion of this program, the GIL has the opportunity to recruit top-of-the line Fellows and contribute to building capacity of young African economists, particularly with respect to gender issues.

GIL’s Effects on Policy

GIL’s research has been used to modify policies, scale-up (or down) innovations, and influence policies indirectly. A conservative staff calculation based on 29 projects is that for every GIL dollar spent, 46 project dollars are being directly influenced.

GIL’s research can be used to modify the design of policies or programs, copy and scale-up (or scale-down) innovations, or influence policy design more indirectly by affecting the policy debate. Evidence of these three different types of GIL research influence to date, gathered through the interviews and review of materials, include:

• Modifying Rwanda’s land registration policy regarding the right of common-law wives (acknowledged in Rwanda’s legal system) to co-own land titles, therefore correcting a flawed design that was having a negative impact on these women’s rights.
• Modifying the operations of one of Kenya’s largest commercial banks by helping set up and test several new forms of joint accounts; modifying the operations of a small private sector bank in Ghana by setting up commitment savings accounts for salaried workers and a pilot program to reach the unbanked.
• Scaling-up the lessons from Liberia’s EPAG in a Republic of Congo youth skills employment program (targeting young women, increasing spaces for them in the entrepreneurship track, and adding financial support for mothers).
• Scaling-down job skills training and condensing business skills training in EPAG’s new training rounds, to make the project more affordable and sustainable.
• Scaling-up the provision of information on earnings on various trades in the Republic of Congo youth skills program based on research in Uganda that found that the provision of this information could increase the probability that women entrepreneurs would choose more profitable (male dominated) trades.
• Copying a personal initiative training tested in a Togo managerial training project in projects in Mexico and Ethiopia.
• Influencing the decision to randomly select the wife who will receive the cash transfer in polygamous households in a Niger social protection program, since a GIL qualitative study of resource and risk sharing in polygamous households in Burkina Faso showed
that adult members in these households maintained individual control of income and assets.

- Through the ‘Leveling the Field’ report, influencing USAID’s internal advocacy and commitment to impact evaluation and evidence-based gender policy making.

GIL staff calculates that they are directly influencing project components totaling US $603 million in 29 projects (86 percent are World Bank operations). According to these staff numbers, for every GIL dollar spent, 46 project dollars are being directly influenced. GIL has not been able to assign a monetary value to at least ten other projects it has influenced directly, as well as the more indirect policy influence it has had.

*Project and data processing delays have slowed down the scope of GIL’s policy influence. GIL could ensure more regular policy influence capacity by staggering evaluations and more fully utilizing results of baselines and midlines. It should also speed up posting of data sets for open access online.*

Much of GIL’s potential for operational influence has yet to materialize, since the majority of the projects in the GIL portfolio are ongoing and have not produced final results. Project implementation delays are common (especially in large World Bank or USAID operations), affecting impact evaluation schedules. In a few cases, projects have ended but GIL has been slow to process and produce final results, and in others, implementers felt that receiving the baseline report in a more timely fashion would have influenced the project positively. There has been frustration with the slowness in producing results, particularly from GIL staff in charge of dissemination and uptake. Two options to better distribute the research workload and shorten the time lag in delivering results are to more systematically stagger the launch of new evaluations and more deliberately utilize the results of baselines and midlines to affect policy. In addition, judicious dissemination of early findings and explanations of the innovation the project is trying to incorporate can be useful. The team has done some of this work through blogs and brownbag sessions.

A good example of using intermediate results to influence policy comes from GIL’s work on the Ethiopia Rural Capacity Building Project (RCBP). The GIL team used early evaluation data to complete an analysis of barriers women faced in accessing agricultural extension. The team, along with the government implementation team, then supplemented the findings with consultations with female farmers to deepen understanding of specific barriers (e.g., not only was the overall amount of time a constraint to access extension, but also specific times during the day). This work helped inspire the government to pilot a special female targeted innovation fund, which was then added to the impact evaluation.

Opening GIL databases has proven to be a time-consuming process. The team has progressed with getting datasets online since hiring an intern to focus on this. The GIL team should begin the process of opening the datasets immediately after a project is closed in order to expedite making the data available, before staff who grant approval for this have moved on to other projects. Once more of the datasets are available online, the GIL may want to consider publicizing the fact that the data sets are available online.

*GIL can increase the likelihood of affecting policy directly by getting in early, evaluating a project’s pilot phases, and having GIL staff participate on project teams. GIL should spot talent for potential GIL staff with combined research and operational expertise and the global practices should offer these staff joint appointments in the practices and the lab.*
GIL’s capacity to affect policies and programs directly is done through the retail support it provides to operations using (a) the project impact evaluation work, (b) direct technical assistance to projects, and (c) GIL staff’s integration into project teams to support the design of Bank operations.

A documented example of (a) above is the change in Rwanda’s land title registration policies already mentioned. Timing or placing the evaluation research around the project’s pilot or initial phase, as was done in Rwanda, increases the likelihood of affecting policy directly.

Instances of one-time direct TA to Bank operations (b) include the qualitative study (rather atypical of GIL) of polygamous households in Burkina Faso for use in the design of a social protection cash-transfer program (to identify who in the household should receive the transfer).

Lastly, being part of a project team (c) is a time-intensive but fail-proof way to embed GIL’s lessons into operations. Currently, only one senior level GIL staff (the thematic leader for entrepreneurship) shares his time equally between research and operations in private sector development. The Bank’s YP program is one of many sources to tap talent and nurture staff that could be shared by GIL and the global practices and that combine expertise in research and operations. This would benefit both the GIL and the global practices, and the Bank headcount policies should be adapted to incentivize these joint appointments.

_A question is if GIL, DEC or the Gender CCSA should evaluate the gender impact of economy-wide reforms, with a view toward influencing policy._

The RCTs that prevail in GIL’s project portfolio can be powerful tools to figure out if project interventions work but have their limitations, particularly in terms of assessing the gender differential effects of and influencing economy-wide reforms. There is some Bank research on these topics, but it is not systematic. A question is whether GIL, the research department or the Gender CCSA should systematically evaluate the impact of economy-wide reforms on women with a view towards influencing their design.

**Promoting Research Uptake**

_The ‘Leveling the Field’ synthesis report combined all of the features of good evidence for policymaking, which helps explain its success in research uptake._

Good evidence for policymaking needs to be unbiased, rigorous, substantive, relevant, timely, actionable, cumulative, and easy to understand and explain. The ‘Leveling the Field’ synthesis report has all of these features, which helps explain its success in research uptake. This report was rated highest among the different types of GIL reports in the client survey.

‘Leveling the Field’ combined rigorous impact evaluation findings with an analysis of the LSMS-ISA, a new household survey instrument incorporating detailed questions on women and gender. The survey analysis framed and increased the generalizability of the evaluation findings. A unique collaboration between the GIL and the ONE Campaign, an advocacy group, which challenged both teams and entailed significant time and work, ensured that the results were easy to understand and explain, and that they would be broadly disseminated. An article on the report

---

in *The Guardian* was picked up by numerous media outlets as was an op-ed by Melinda Gates in Project Syndicate.

‘Leveling the Field’ shows that successful research uptake requires a substantive, rigorous and actionable message combined with expert communication skills and, if possible, independent high-level endorsement. For maximum research uptake, partnering with groups who do not ‘speak the same language’ around impact evaluations may be critical, in order to identify areas that may not be clear to policymakers and break down silos between sectors.

*Aside from the ‘Leveling the Field’ report, GIL has only sporadically promoted the process of research uptake, partly because a critical mass of results is not yet available.*

GIL designed a detailed Research Uptake Strategy, and put in place a dissemination structure (with web sites, materials with designs, policy blogs, etc.), all perhaps too early for the slower pace at which results have trickled in. The pace of results should speed up significantly next year, and to ensure that the process of uptake functions well, GIL needs to reinstate a senior staff position for research uptake that can interface with both operational teams at the Bank and external stakeholders. This position could become part of the proposed federation of GILs and be shared with other regional GILs.

GIL researchers can also facilitate research uptake by taking proactive steps to meet with specific individuals and agencies in country. There is an in-country checklist, part of the Research Uptake Strategy, which could be revisited by the full team to bolster outreach efforts. Effort should also be made when in-country to reach out to other important stakeholders, including the international agencies’ gender group and GIL donors’ offices.

*Dissemination of GIL research and activities within and outside the Bank has progressed over time.*

Examples of Bank-wide dissemination are the GIL’s biweekly seminar series, and a well-attended Brown Bag Lunch session called a ‘smack down,’ which is organized approximately every six months and brings together researchers and practitioners for debate. Attending staff meetings is also perceived as a good way to get the message out about the GIL work, but follow up and meeting requests from operational teams can be challenging due to time constraints. The GIL team is now asked regularly by operational teams for technical input on their projects. As mentioned above, to strengthen further the link to operations, GIL should consider the possibility of having more staff that share time with operations.

Sixty-six percent of respondents to the client survey agreed that the GIL somewhat effectively or effectively made its results available to a broad audience. The most useful communication vehicle was the synthesis report, followed by working papers and policy notes. As of mid-2015, GIL had published 10 policy briefs (compared to 15 originally targeted), 8 working papers (5 originally targeted), and 1 synthesis report (0 originally targeted). This year alone, the GIL has presented its work more than 75 times, which far exceeds its original target.

**GIL’s Organizational Structure, Management, and Budget**

*GIL is ‘on track’ to meet or exceed the goals set in the proposal to DFID.*

According to DFID’s funding cycle, to be considered ‘on track’ there should be 18 GIL 1.0 and 12 GIL 2.0 endline surveys completed by the end of 2017, and the GIL is on track to meet or
exceed these goals. Progress toward current targets and indicators was recently reviewed by the GIL team for the 2015 annual report, and is captured in Table 2 of this report.

**Administering the GIL from the UFGE works well and helps to achieve the GIL’s objectives, including around research uptake.**

Benefits of administering the GIL from the UFGE include cross-fertilization and learning between regions and therefore amplifying the impact of the GIL across the Bank. For example, the African Region GIL has been used as a model for setting up EAP and LAC Labs’ administration procedures. Going forward, being under one umbrella will likely facilitate strategic alignment of the Labs as part of the Bank’s overall (forthcoming) gender strategy. The planned model, a federation of loosely affiliated labs all sitting under the Gender CCSA, is likely to allow learning across regions with respect to administration procedures and technical innovations. This would need to be supported by dedicated staff time. The UFGE could play a stronger role in fostering innovation by fundraising for the Labs across regions, perhaps for comparative work, or to fill gaps not prioritized by regional funders. The UFGE or the proposed federation should also consider adopting an ethical protocol for research with human subjects.

The UFGE has provided a large degree of independence to the GIL to administer its own grants, which has allowed for flexibility to respond to regional needs and opportunities. This element is critical to the success of the GIL and should not be changed. At present, both the GIL and the UFGE have their own direct relationships with DFID, which works well. DFID provides valued technical input to the GIL.

**The GIL team has a flat reporting structure that has evolved successfully along with GIL 1.0 and 2.0 structures, but it may be wise to establish mid-level operational and technical management as the team grows in size.**

GIL staff hiring has been done with careful attention to ‘fit,’ and this has contributed significantly to GIL’s overall success. Team members are self-motivated and highly committed to the work, allowing a more flexible management structure to be successful. With no direct line manager, team members commonly review each other’s work, and this process is very collaborative. On its own initiative, GIL recently administered a management survey to the full team. Results were positive overall and management is taking steps to address concerns raised, which are related to staff receiving adequate feedback and mentoring. As the team expands in size, it may be wise to put in place a mid-level operational and/or technical manager so that adequate feedback and learning are possible.

**Adequate structures are in place to ensure the GIL discontinues projects that are not going forward as planned after several attempts.**

A total of five projects have been dropped from the GIL. These were dropped because implementation would be delayed beyond 2017, or, upon exploring further, it was clear that it would not be possible to get an adequate sample size.

When it is unclear whether a project will come to fruition, or it seems it is taking too long or becoming too expensive, the GIL team sets benchmarks for its project coordinators, and this motivational system appears to work well to keep most projects on track. When benchmarks are not met, the projects may be dropped.
A more systematized method of capturing research uptake would be useful, particularly as research results are made available.

The system for capturing research uptake activities and their resulting outcomes is somewhat ad hoc, and several GIL team members expressed an interest in formalizing it. At present, one GIL staff member requests information on dissemination activities from the rest of the team on a monthly basis, and codes this information for reporting purposes.

Following this mid-term review that attempts to more precisely define research uptake, it may be valuable for the GIL team to devise a common understanding of research uptake and the types of activities that should be reported on. It may be helpful for the team to review the coding system together so that all members are aware of how research uptake is being defined and captured through both direct and indirect activities.\textsuperscript{12}

The GIL budget is on track and is managed closely and competently.

The GIL budget has benefited from close management and administration by a staff member who has set up a system of incentives to ensure timeliness and adherence to the allocated funding amounts. The funding is released in tranches at key project milestones.

Per project costs are on par with standard impact evaluation costs, and are targeted at $750,000 per evaluation, with fairly significant variation between projects.

The table below shows a breakdown of spending for GIL 2.0 across key categories since 2013.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
\textbf{Bank Budget} & 2013-2014 & 2015 & Total \\
\hline
Various (impact evaluations, staff costs, RA costs and travel) & 1,600,000 & 1,500,000 & 3,100,000 \\
\hline
Trust Fund & - & - & - \\
Staff costs & 1,890,000 & 1,625,000 & 3,515,000 \\
Impact evaluation costs & - & - & - \\
1.0 & 828,000 & 792,000 & 1,620,000 \\
2.0 & 1,980,721 & 4,632,027 & 6,612,748 \\
Capacity Building & 95,658 & 129,342 & 225,000 \\
Research Uptake & 10,100 & 39,900 & 50,000 \\
Governance and other M&E & 13,700 & 26,300 & 40,000 \\
\hline
Total & 6,418,179 & 8,744,569 & 15,162,748 \\
\hline
\end{tabular}
\end{table}

\textsuperscript{12} Motivated by this mid-term review, GIL has already made significant headway in better capturing research uptake, and provided useful information for the review.
Conclusions and Suggestions Looking Forward

It is clear from this mid-term review that the Africa GIL is a very high quality, well-run evaluation research initiative that has already contributed rigorous evidence on what works to increase women’s productivity and income, and influenced development policies directly and indirectly. We agree with an independent mid-term review of the UFGE that states that the Africa GIL is a particularly successful strategic grant, which adds value and complements the Bank’s gender mainstreaming work by playing a catalytic role in the generation of new ideas with significance beyond the Africa region.¹³

The GIL has four interdependent activities: evaluation research, workshops and fellowships, retail support to projects, and publications and communications, all essential to promote the process of research uptake and influence policy. The GIL should strengthen its Fellows program, an important component of expanding policy capacities in Africa. GIL can also do more to promote research uptake more systematically or regularly, even when most of the projects in the portfolio are ongoing and have yet to produce final results. It would be smart to strengthen these GIL functions now, before the expected onslaught of results in 2017.

The different sections of this review contain a number of suggestions going forward for the different categories of GIL activities, from exploring more in depth how heterogeneity among women affects outcomes, to staggering impact evaluations (or making two calls at separate times), to reinstating a senior staff position for research uptake. Most assume some (moderate) growth in GIL’s budget and at very least that GIL will continue to operate at current funding levels.

The Bank’s gender strategy, under preparation, suggests, however, an alternative scenario of greater growth for the Africa GIL, as a lead GIL in a federation of regional GILs with a broader mandate that includes diagnostic as well as evaluation work. The Bank needs to reflect carefully on the relative merits of this expansion. While the GIL has the expertise and the tools for deep-dive country diagnostic work, this work is an essential component of gender mainstreaming at the country level that the Bank global practices should undertake. The uniqueness of the GIL is in researching what works to promote women’s economic empowerment, a ‘public good’ type activity that is otherwise underfunded at the Bank and externally. Our preference would be for the GIL to continue to focus on ‘What Works’ and excel at evaluation research, and strengthen its interdependent activities to maximize research uptake and policy influence through planned, gradual growth.

### Table 1: Sample of GIL’s Impact Evaluations

See accompanying Excel sheet.

### Table 2: Progress Toward Meeting Quantitative Targets

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Target for 2015</th>
<th>Progress Made by 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of scaled up effective gender-related policies and programs</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Number of scaled down or avoided ineffective gender-related policies and programs</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Presentation of results to decision makers within a development constituency</td>
<td>50</td>
<td>76</td>
</tr>
<tr>
<td>Number of academic papers and policy reports citing gender lab studies</td>
<td>25</td>
<td>135</td>
</tr>
<tr>
<td>Number of new rigorous impact evaluations commissioned through a transparent selection process (i.e. concept notes approved)</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>Number of impact evaluations (ongoing and new) with completed endline surveys</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>Number of impact evaluations (ongoing and new) that measure agency and empowerment impacts</td>
<td>23</td>
<td>36</td>
</tr>
<tr>
<td>Number of policymakers that take part in workshops and report increased knowledge and understanding of impact evaluations or findings</td>
<td>120</td>
<td>270 participants, feedback from 69</td>
</tr>
<tr>
<td>Number of (male/female) Junior and Senior Fellows</td>
<td>9</td>
<td>2, both male</td>
</tr>
<tr>
<td>Number of peer reviewed publications with a Southern researcher as a primary author</td>
<td>2</td>
<td>5 published</td>
</tr>
<tr>
<td>Number of published research outputs, including synthesis papers, policy briefs and working papers from impact evaluations published and available freely online. (Number of peer reviewed publications)</td>
<td>15 policy briefs; 5 working papers (3 peer-reviewed); 0 synthesis papers</td>
<td>10 policy briefs; 8 working papers (5 peer reviewed); 1 synthesis paper</td>
</tr>
<tr>
<td>Number of data sets made openly and freely available to external researchers</td>
<td>30</td>
<td>8</td>
</tr>
<tr>
<td>Development and effective implementation of research uptake strategy</td>
<td>1</td>
<td>1 developed, implementation ongoing</td>
</tr>
</tbody>
</table>

### Table 3: List of People Interviewed

**GIL team group interview**
Salman Alibhai, Gautam Bastian, Ioana Botea, Niklas Buehren, Francisco Campos, Eliana Carranza, Rachel Coleman, Aletheia Donald, Brittany Hill, Tigist Ketema, Edouard Mensah, Michael O'Sullivan, Sreelakshmi Papineni, Lea Rouanet, Nenette Santero, Julia Vaillant

**GIL team individual interviews**
Francisco Campos
Eliana Carranza
Shubha Chakravarty
Rachel Coleman
Malcolm Ehrenpreis
Markus Goldstein
Brittany Hill
Michael O’Sullivan
Rachael Pierotti

**Junior Fellows**
Zewdu Ayalew Abro,
Edouard Romeo Mensah

**Development partners**
Vanessa Adams
Juliette Seban
Wessel van den Berg

**DFID**
Lina Payne

**USAID**
Jessica Cagley
Bhavani Pathak
Farzana Ramzan

**Government partners**
Eusebe Agoua
Michael Kouba

**GIL Steering Committee**
Elizabeth King

**Internal World Bank clients**
Eneida Fernandes
Victoria Monchuk
Francisco Strobbe

**Other World Bank staff**
Makhtar Diop
David Evans
Caren Grown
Sarah Nedolast
Sudhir Shetty
Dominique van de Walle