Motivation

Take advantage of scope and scale of DIME research to develop and test econometric and technical solutions

Improve quality of data collection and research methods across DIME’s IE portfolio

Make public trainings and standardized software tools available to the community of development researchers
A systematic, iterative process to expand technical capacity across the research cycle

- Conceptualize improved workflows
- Develop new tools and processes
- Institutionalize practices through training and support
- Identify inefficiencies during project support

Reproducible Research
Quality Assurance for Data Collection
Public Resources and Training
Open-source Research Tools
Reproducibility
Reproducibility at DIME

• New **DIME Research Standards** (see handout for discussion)
  – Research Ethics
  – Research Transparency
  – Research Reproducibility
  – Data Security
  – Data Publication

• Aim for 100% adoption of standards across portfolio
Reproducibility at DIME

• Planned first push: “Bootcamp” November 11-13
  – 2 days of mandatory hands-on training for all RAs
    • Objective: each RA implements all applicable research standards for one project (e.g. encryption, secure data storage, code on GitHub, etc)
  – ½ day training with TTLs + RAs to discuss project updates and recommended work flows (e.g. project management in github)
    • Enough? Suggestions on what is needed to effectively supervise welcome.

• Analytics will send out a survey this week, which will:
  – map RAs to projects (to understand how much of portfolio we’ll cover)
  – prioritize projects for the bootcamp
  – check availability the week of November 11th (can schedule second date if needed)

• We will reach out to project TTLs before bootcamp regarding team composition and preferences on data organization
Pre-publication code review

• Analytics reviews the code for all DIME working papers to ensure replicability before publication
  – To add to paper acknowledgements: “Computational reproducibility verified by DIME Analytics. Please see details of the reproducibility check in the online appendix.”

• Completed reproducibility checklist will be returned to the team to add to online appendix
Peer code review

• 2 days per year (1 every 6 months)

• Mandatory participation of all RAs – these days will be completed blocked

• Opportunity to review intermediate outputs (e.g. IE reports)

• Analytics will provide feedback to TTLs on participation and strengths / weaknesses of code reviewed
Transparency, Reproducibility & Credibility: A Research Symposium

- Collaboration with 3ie, BITSS, IPA, and J-PAL

- Symposium materials
  - Presentations
  - Video of DIME Analytics lightning talk
  - Blog: What development economists talk about when they talk about reproducibility
  - Data and results from Reproducible Research survey by Lupe and team

10/9/19
Quality Assurance for Data Collection
DIME Survey Library

• Current situation:
  – 265 + surveys
  – Questionnaires live on hard drives: no centralized access, and no way to recover questionnaires when staff leave
  – RAs and FCs tasked with survey design reinvent the wheel

• Where we want to be:
  – Centralized questionnaire repository (ideally taggable and searchable)

• What’s next?
  – Compile questionnaires and relevant documentation – look for a request linked with the myIE update
Survey Review

• **Objective**: improve and standardize quality of field data collection, ensure field-based staff are aware of all DIME resources and protocols

• **On-demand service**: email dimeanalytics@worldbank.org to set up initial consultation (as early in survey planning process as possible)

• **Customized support**: depending on needs of team

<table>
<thead>
<tr>
<th>Field plan</th>
<th>Research ethics and data security</th>
<th>Survey firm procurement</th>
<th>Questionnaire review</th>
<th>Sampling</th>
<th>Data quality assurance plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Review timeline</td>
<td>• References on local IRBs</td>
<td>• Template TORs</td>
<td>• Check questionnaire for best practice &amp; organization</td>
<td>• Technical assistance with sample design and power calculations</td>
<td>• Guidelines for data quality assurance plan</td>
</tr>
<tr>
<td>• Technical assistance on survey organization</td>
<td>• Boilerplate text for informed consent</td>
<td>• Technical assistance for procurement questions</td>
<td>• Test CAPI programming</td>
<td></td>
<td>• Review and feedback of plan</td>
</tr>
<tr>
<td></td>
<td>• Review of data security protocols</td>
<td>• Participation in selection committee for DIME procurements</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Instructions for data encryption</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

10/9/19
SurveyCTO and Survey Solutions

• SurveyCTO server currently funded through January 2020
  – Expect it to continue through January 2021

• Exploring transition to Survey Solutions for new projects

• Planning pilot: side-by-side comparison of the tools with two DIME surveys over the next 6 months, to confirm that Survey Solutions meets all DIME’s needs
Public Resources and Training
Data for Development Impact

- A new “user manual” for development data
  - Complements the DIME Wiki with a narrative approach
  - Designed for training new RAs or to use as a desk reference

- First version available. Revising this fall (feedback welcome!). Updated version in early 2020.
Manage Successful Impact Evaluations

• Flagship training, held annually

• Designed to improve skills & knowledge of IE implementers, e.g.:
  – Design and program electronic survey instruments
  – Monitor data quality and provide real-time feedback
  – Manage survey data and produce descriptive analysis for policymakers

• MSIE 2019 agenda with videos, and all training materials

• 2020 dates: June 15-19
DIME Wiki

• One-stop shop for impact evaluation research solutions, launched 2018

• Global audience & user-base
  – 239 registered contributors
  – Viewers from 197 countries
  – 8 most active cities after DC
    • Nairobi, Chicago, New York, Kampala, London, Lagos, Dar-es-Salaam, Bengaluru

• Continuing to build content; linkages to D4DI
Build in-house capacity

• Expanding Continuing Education to Field Coordinators (1 session per month + 1 open office hours per month
  – Starting October 24th

• On-demand sessions for RAs (content as needed following up RA retreat)

• Interest inside the team for a new round of R training?
Open-source Research Tools
Open-source Research Tools

Standardize and simplify best practices for reproducible IE data

- iefieldkit: Tools for primary data collection and processing
- ietoolkit: Tools for data management and analysis
- visual library: Reproducible code for graphs and maps in an easy-to-browse format
**iefieldkit**

- Suite of Stata commands to routinize common tasks for primary data collection
  - `ietestform`
  - `ieduplicates & iecompdup`
  - `iecodebook`

- Article submitted to Stata Journal

- `iecorrect`: new command to make HFC corrections
ietoolkit

- Stata commands for data management
  - `ieboilstart`
  - `iefolder`
  - `iegitaddmd`
  - `ieboilsave`

- and IE analysis
  - `iebaltab` and `ieddtab`
  - `iematch`
  - `iegraph`

- `iekdensity`: new command to display variable distribution over multiple groups
- `iecompressall`: new command to reduce folder size
- New command to check push-button replication?
Visual Library

Stata Visual Library

Explore the library
This page contains graphs created in Stata to explore impact evaluation data sets and display results. Click on a graph to view the code that created it.

GitHub Repository
The do files and data sets used to create the Library’s graphs are stored in the Stata Visual Library GitHub repository.

Useful links
- DIME Wiki
- Stata Graphics
- Stata Cheat Sheets
- ietoolkit
- Stata GitHub

Adding content and search function
Tool to review .do files

• Small app that will flag do files for stylistic and construction errors.

• Functions as the first step in code review process

• RAs can use the tool at their convenience to make sure that the code follows the guidelines prescribed by DIME Analytics.
RA & FC Support
RA & FC Support

• New hire form
• RA and FC onboarding
• Software installation
• Office hours (request form)
• DIME Alumni listserv (join here)
• Drafting data use / non-disclosure agreement for RAs to sign at onboarding
FC Onboarding

• If you fill out this survey when new FCs are hired …
  – FCSupport will then send an onboarding guide to your consultants, connecting them with all DIME Analytics resources
  – FCSupport will schedule a virtual FC onboarding session

• When you hire a new FC:
  – Make sure to inform the Country Office (typically through SMO)
  – Make sure to inform the Global Security Focal point at the CO
FCSupport Team

• Create and manage a database of all consultants actively working for DIME
  – NB. If you haven’t sent back your list of FCs yet, please do it now!

• Provide first point of contact support for FCs (fcsupport@worldbank.org)
  – Re-direct FCs toward the relevant person/resource when they have questions and do not know who to contact
  – Connect them with senior field coordinators working on similar countries or topics
Analytics Team

Maria Jones  |  Luiza Andrade  |  Kristoffer Bjarkefur  |  Benjamin Daniels  |  Judith Indalecio

Roshni Khincha  |  Mrijan Rimal
Thank you!