

Public Disclosure Authorized

r-tm

Measuring welfare when it matters most

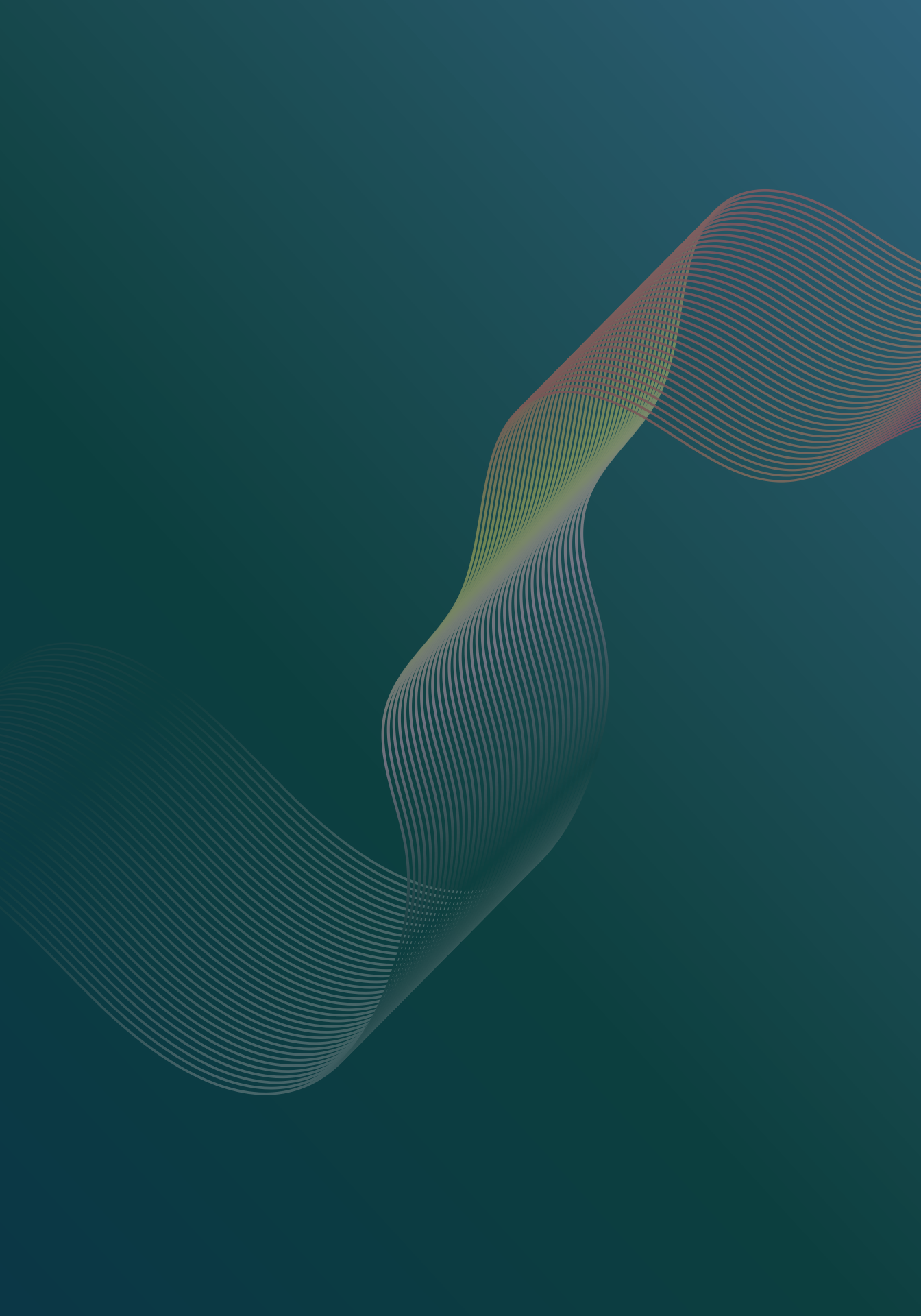
Public Disclosure Authorized

Public Disclosure Authorized

Public Disclosure Authorized



A typology of approaches for real-time monitoring



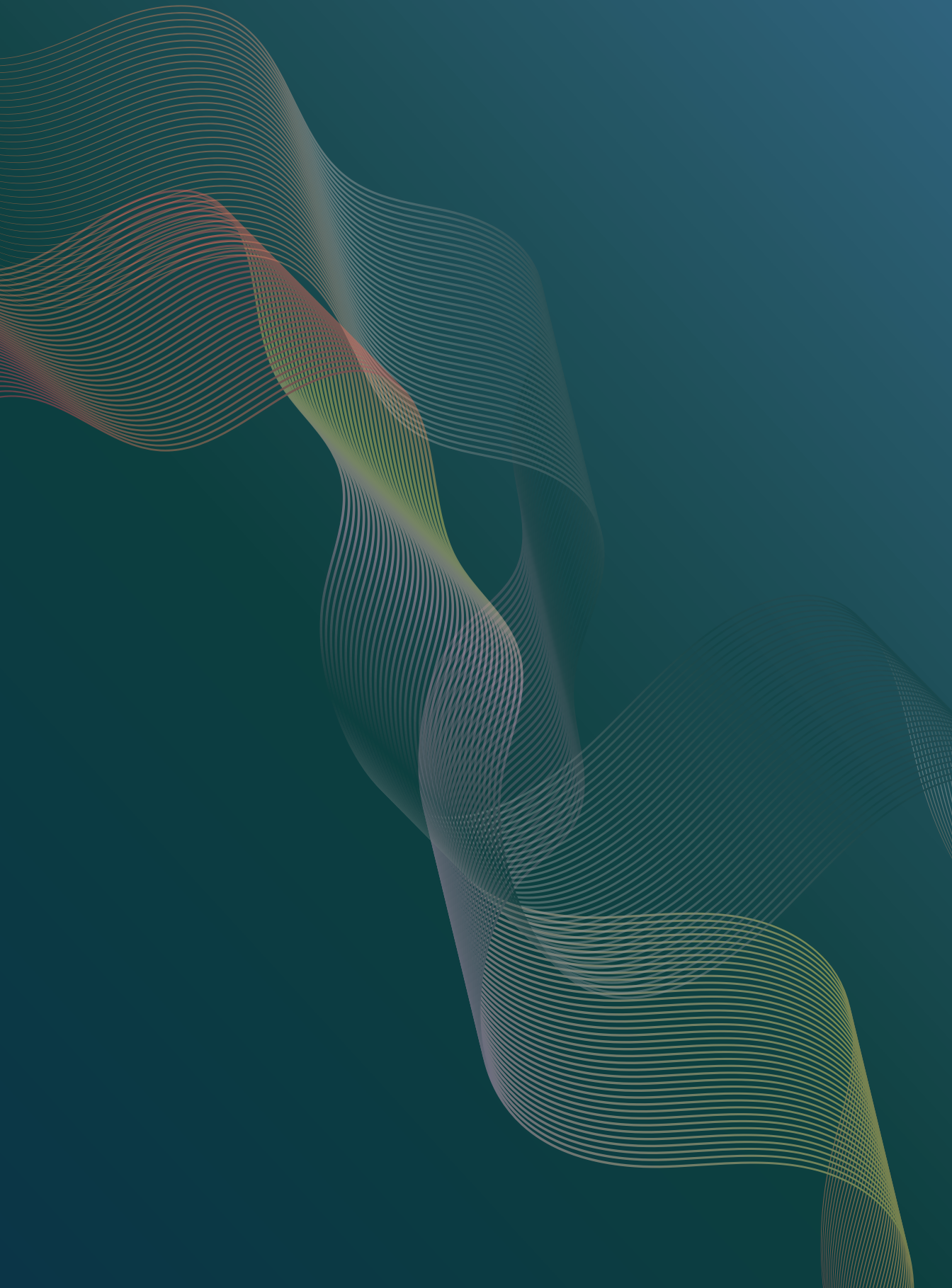
Contents

| | |
|--|-----------|
| Introduction | 7 |
| 1 Methods for Nowcasting Welfare—With a Focus on Monetary Poverty | 17 |
| 1.1 Nowcasting Welfare Using Survey and Other Non-survey Covariates | 18 |
| Considerations Regarding Reliable Survey- and Non-survey-based Imputation | 21 |
| Survey Imputation Methods can be Complemented with Data Collection to Deal with Missing Auxiliary or Baseline Data | 24 |
| Lessons Learnt and Resources | 25 |
| 1.2 Nowcasting Welfare Using GDP Growth | 27 |
| Considerations Regarding GDP-based Nowcasting | 30 |
| Resources | 33 |
| 1.3 Nowcasting Welfare Using Microsimulations and General Equilibrium Models | 33 |
| Considerations Regarding Microsimulation and General Equilibrium Models | 36 |
| Resources | 37 |
| 2 Harnessing Data for Real-time Welfare Monitoring | 39 |
| 2.1 Rapid Survey Data Collection | 40 |
| High-frequency Phone Surveys | 40 |
| Rapid Face-to-face Surveys | 46 |
| Online and Messaging-based Surveys | 49 |
| Further Resources | 50 |
| 2.2 Geospatial Data | 52 |
| Main Characteristics and Examples | 52 |
| Caveats for Using Geospatial Data | 56 |
| Lessons Learnt and Resources | 59 |
| 2.3 Digital Trace Data | 62 |
| Main Characteristics and Examples | 62 |
| Caveats for Using Digital Trace Data | 64 |
| Lessons Learnt and Resources | 65 |

| | |
|---|------------|
| 2.4 Administrative Data | 67 |
| Main Characteristics and Examples | 67 |
| Caveats for Administrative Data for Real-time Welfare Monitoring | 69 |
| Lessons Learnt and Resources | 70 |
| 3 Moving Forward: Identifying Areas for Advancement | 71 |
| References | 73 |
| Annex 1. Summary of Models Used to Update Poverty Estimates | 95 |
| Annex 2. Commonly Used Machine Learning (ML) Models for Estimating Poverty | 97 |
| Annex 3. Summary of All Data Sources | 100 |
| Annex 4. Nowcasting Impacts of Shocks (Vulnerability and Damage Functions) | 103 |
| Considerations Regarding Damage Functions | 104 |
| Resources | 105 |

Acknowledgments

This draft was prepared by a team from the World Bank Poverty and Equity Global Practice consisting of Kimberly Bolch, Maria Eugenia Genoni, and Henry Stemmler. Carlos Sabatino also provided excellent inputs to the document. The work was conducted under the supervision of Luis F. López-Calva (Global Director, Poverty and Equity GP) and Benu Bidani (Practice Manager, Poverty and Equity GP). This document benefited from consultations with many members of the Poverty and Equity Global Practice as well as other World Bank teams who led the development and implementation of many of the initiatives referenced here. The team is particularly grateful to Maurizio Bussolo, Paul Corral, Ximena Del Carpio, Daniel Gerszon Mahler, Craig Hammer, Ruth Hill, Dean Jolliffe, Walker Kosmidou-Bradley, Laura Moreno Herrera, Sergio Olivieri, and Nobuo Yoshida for their comments and advice. Design and typesetting by Reyes Work.



Introduction

Timely Information on Welfare is Critical for Effective Policymaking

As the *World Development Report 2021: Data for Better Lives* highlights, data is a foundational input for improving development outcomes through enhancing the effectiveness of policymaking. However, the degree to which data can generate value for development depends on its quality (World Bank, 2021b). One critical aspect of data quality is timeliness. Having up-to-date information is essential for policymakers to finetune policies as conditions change. In a contemporary global environment marked by heightened uncertainty in the face of challenges such as climate change, conflict, and pandemics—the need for more timely sources of data to inform policy is particularly pressing.

In the context of policies to reduce poverty and vulnerability, more timely information on household welfare is needed. Traditional methods produce measures of household welfare too infrequently to meet the needs of many policymakers. Official measures of poverty are derived from household surveys, which (even in ideal settings) are only conducted every few years—given the financial and administrative costs involved. In many settings, and particularly in low-income and fragile countries, these surveys are conducted with much greater lags.¹ However, by combining traditional surveys (“baseline data”) with different modelling approaches and alternative sources of frequently collected data (“auxiliary data”)—it is possible to develop monitoring systems that

¹ On average, the most recent household survey in the World Bank’s [Poverty and Inequality Platform](#) (PIP) is over six years old. Of the 168 countries in PIP, 37 percent have data that is more than five years out of date; of the 56 IDA countries in PIP, 52 percent have data that is more than five years out of date (September 2023 PIP Update).

provide up-to-date estimates on the evolution and status of household welfare. Investing in this capacity to monitor welfare in “real time” is essential to both (i) inform new policy action in the wake of shocks and (ii) enhance the adaptive capacity of existing policies as circumstances change. In addition to serving as inputs to effective policymaking, many of the approaches discussed can be applied in the context of project monitoring. See Box 1 for a discussion on how we define monitoring of "welfare" in "real time".

Methodological and technological advances have expanded our ability to monitor welfare in real time

In recent years, the World Bank’s Poverty and Equity Global Practice (GP) has increased its capacity to provide more timely information on welfare. In close collaboration with internal and external partners, we have led efforts at the country (responding to context specific needs) and corporate levels (related to the global monitoring of poverty) to implement a broad range of modelling approaches and leverage or collect new sources of high-frequency auxiliary data. Moreover, this work has increasingly benefited from frontier methodological approaches (for example, machine learning) and data sources (for example, big data) that can enhance the performance of existing methods. While ongoing for some time, the work was greatly scaled up in the context of recent crises such as the COVID-19 pandemic and climate-related disasters.

This typology takes stock of the growing body of work on real-time welfare monitoring, bringing together existing resources and lessons learned in one place. It aims to offer an overarching roadmap to help teams navigate different approaches and identify the best fit for answering a specific question in a given context. The “best fit” approach may differ across settings depending on a country’s data ecosystem and implementation constraints. This typology systematizes the decision-making process by laying out the various advantages, disadvantages, underlying data requirements, and assumptions of different approaches. While primarily drawing the Poverty and Equity GP’s work, the typology aims to contextualize real-time monitoring within a broader body of research and towards recent innovations in the field. The research this typology has produced is part of a broader global initiative of the GP on moving “Towards Real-Time Monitoring of Welfare” and will be complemented by a more detailed technical handbook (forthcoming).

Box 1 Defining Real-time Welfare Monitoring

What do we mean by “real time”? This typology uses the term “real time” to refer to information produced with a shorter lag than traditional household surveys allow. For welfare monitoring, where survey gaps often span multiple years, data produced with weekly, monthly, or even yearly periodicity may be considered “real time.” The goal of the various approaches described in this typology is to provide the most up-to-date welfare information possible, given the feasibility constraints for doing so reliably. It does not necessarily imply instantaneous updates.

How do we define “welfare”? We use the term “welfare” broadly to encompass multiple dimensions of well-being. The typology highlights examples of welfare monitoring across a range of dimensions with a focus on monitoring monetary poverty at the national level, reflecting the extensive work produced by the Poverty and Equity GP on this aspect of well-being.

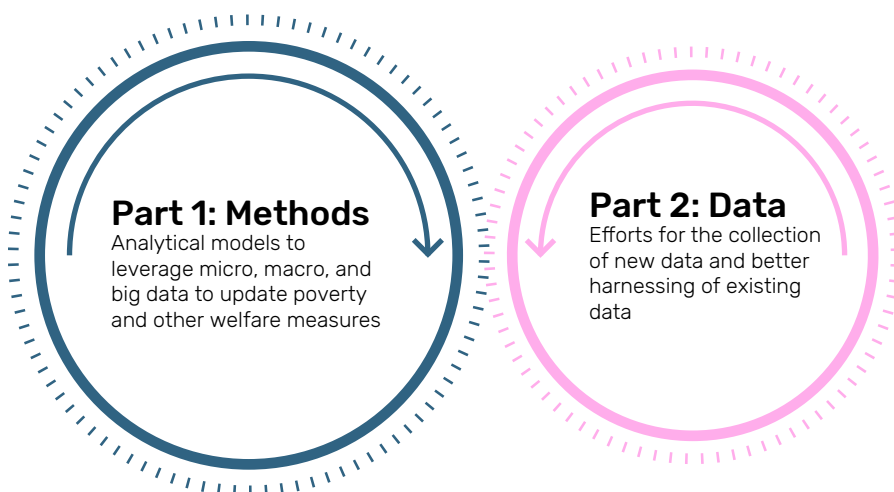
Monetary poverty is a state of deprivation characterized by a lack of sufficient income or financial resources to meet basic needs, such as food, shelter, clothing, and healthcare. It is typically measured by comparing an individual’s or household’s income or consumption against a defined poverty threshold or poverty line, with those below the threshold considered monetarily poor.

Monetary poverty measurement is data intensive and challenging in data-deprived contexts. In some cases, directly measuring other dimensions of welfare (for example, food security, employment, housing, education) may be easier and equally insightful for understanding changes in individual well-being.

A Typology in Two Parts: Methods and Data

This typology is organized in two parts. The first part focuses on **methods**, mapping out analytical models that leverage micro, macro, and big data to nowcast poverty and other welfare measures. The second part focuses on **data**, listing options to collect high-frequency data or better harness existing sources. Most approaches require a strategic combination of both—with models requiring high-frequency data as a key input (Figure 1).

FIGURE 1 Real-time welfare monitoring requires a combination of modeling and high-frequency data



Notably, most approaches rely on having recent baseline data as a precondition (Box 2). In this sense, these approaches are not meant to be a substitute for investing in traditional surveys (such as household budget surveys or censuses); in fact, having a relatively recent baseline survey is a critical input to ensure the quality and accuracy of the modeling and data collection methods covered in this typology. When this is not the case, the feasibility of real-time monitoring may be limited, and the collection of new baseline data may be required.

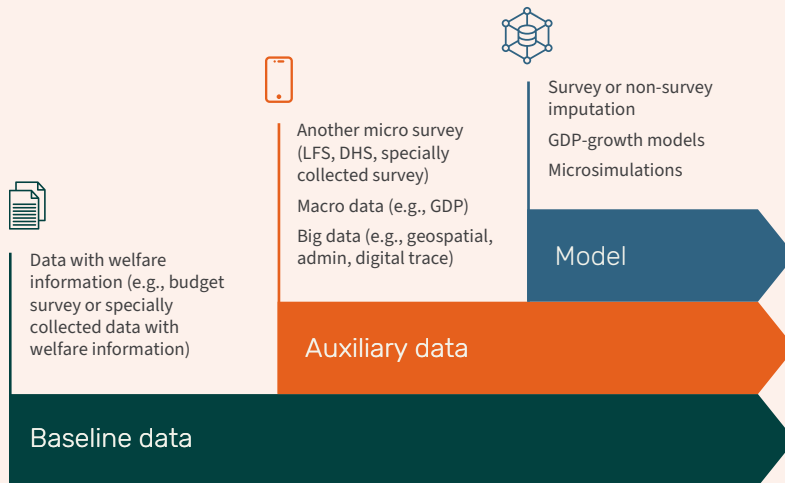
Box 2 Building on a Strong Foundation: Baseline Data is a Prerequisite for Real-time Monitoring

Methods for imputing poverty and high-frequency “auxiliary data” are not yet substitutes for traditional household surveys, which remain the foundation of reliable welfare estimates. A full survey with comprehensive welfare information (such as a household budget survey) or population information (such as a census) is often a prerequisite to effectively apply the approaches described here.

Figure 2 shows how to think about these different data sets and how they together feed into models to monitor welfare in real-time. This typology refers to this foundational data as “baseline data.” However, in the language of machine learning it can also be thought of as “training data.” Training data serves a pivotal role, providing the underlying information necessary for models to learn patterns, classify data, and make predictions. The quality and quantity of training data significantly impacts the performance and accuracy of the algorithm. If training data on welfare is non-existent or too out of date, these methods will be unreliable.

The methods and data sources discussed in this typology should be seen as complementary to rather than substitutes for traditional surveys. As such, efforts to advance the real-time monitoring of welfare greatly depend on continued investments in closing foundational data gaps. The World Bank has long been working with country partners to invest in the modernization of national statistical systems. At the global level, this work is being led by the Global Solutions Group on Data for Policy. This includes an important effort to close poverty-related data gaps, including through the implementation of more frequent household surveys. While much progress has been made in recent years, there is still a long way to go.

FIGURE 2 The ingredients for real-time welfare monitoring



Part I: Methods

This portion of the typology provides an overview of various types of methods that can be used to impute or predict welfare in “real time”. These methods utilize timely information from “auxiliary data” sources (such as micro surveys, macroeconomic statistics, or other big data sources) and model relationships with variables in older baseline data to estimate missing data points.

Figure 3 provides an overview of methods of real-time monitoring for different use cases. The main types of methods discussed in this typology are covariate-based nowcasting, GDP-based nowcasting, and microsimulation models. Researchers have all these methods at their disposal when the objective is to obtain an updated poverty-rate nowcast. GDP-based nowcasting needs to be modified to capture differences across income distribution, while the other methods incorporate distribution sensitive nowcasts. When researchers aim to incorporate different mechanisms and indirect effects, they need to rely on microsimulation models. Finally, microsimulation models and related vulnerability functions are useful for updating estimates to account for the impacts of shocks. Covariate-based nowcasting can also provide estimates of shock

impacts, but typically only when combined with data collection efforts, which are discussed in further detail in part 2 of this typology.

FIGURE 3 Methods for real-time monitoring for different use cases

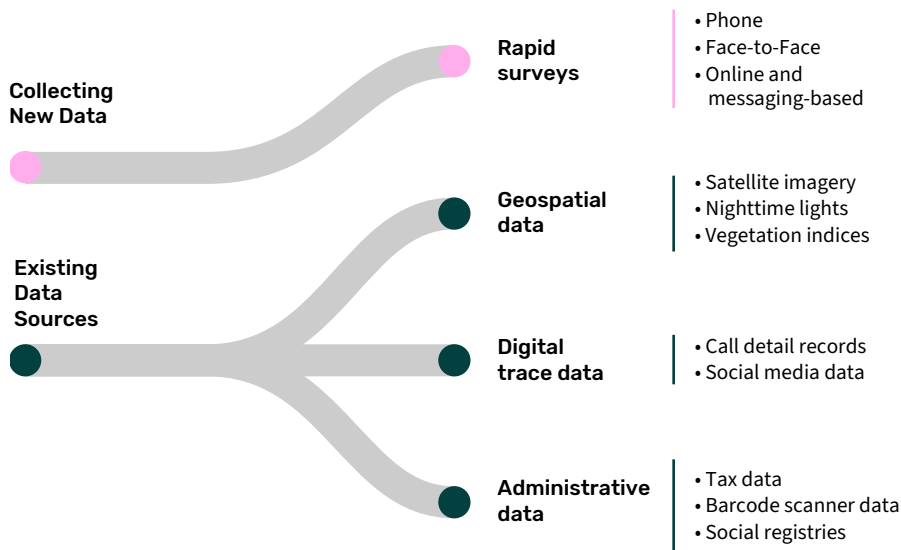
| Use case | Methods | | | |
|---|--|--|--------------------------------------|-------------------------------------|
| Poverty-rate nowcast | Covariate-based nowcasting (section 1.1) | Distribution-scaling (section 1.2) | GDP-poverty elasticity (section 1.2) | Micro-simulation (section 1.3) |
| Estimates along the income distribution | | Can incorporate assumptions about distributional changes | | |
| Incorporate and understand mechanisms or indirect effects | | | | |
| Nowcasting changes in welfare after shocks | Collection of ex-post data (section 2.1) | | | Vulnerability analysis (Appendix 4) |
| Monitor proxy or leading indicators for welfare | Harnessing data (section 2) | | | |

Ultimately, choosing between the various methods will depend on the use case and to a large extent on the underlying data requirements and the scale of analysis (for example, subnational, national, regional, global). Moreover, implementing these methods requires different inputs in terms of skills, time, and financial resources. Depending on the constraints that a team faces in a given context, different approaches may be better suited to the realities on the ground. This typology features several decision trees to help users think through which method(s) are better suited to different contexts and objectives.

Part II: Data

The timeliness of welfare estimates produced by the methods depends entirely on the timeliness of the auxiliary data inputs. Reliable, high-frequency and up-to-date data sources are critical for any approach to monitor welfare in real time. Part II of this typology focuses on two key efforts: (i) collecting new high-frequency data, and (ii) better harnessing existing sources of high-frequency data (Figure 4 illustrates a few examples).

FIGURE 4 Data for real-time monitoring: Collecting and harnessing high-frequency data



These more frontier types of data sources can be leveraged in several ways for real-time monitoring. First, they can complement existing baseline survey data as an input to improve the nowcasting methods discussed above. This can be particularly useful when existing survey data is not recent, does not cover the whole population, or lacks specific dimensions that are relevant for welfare estimation. Second, they can offer a broader picture on welfare when data constraints limit the feasibility of estimating monetary welfare. In many cases, other (non-monetary) measures are very informative in depicting welfare trends or differences between populations. Variables such as employment, food security, or subjective well-being may be available from other sources or can be collected more easily than full information on income or consumption. Third, leading indicators, such as predictions of droughts or floods or inflation data, can provide important signals of changes in welfare, before these are observable in survey data. These last two use cases are summarized by the last row of Figure 3.

Selecting the best-fit approach

This typology is not meant to be prescriptive nor does it rank approaches. Rather, it seeks to provide a more structured way to help users identify a core set of available options and systematically think through the trade-offs

between them. Each approach covered in the typology includes a discussion on the main characteristics, caveats, and lessons learned—alongside a collection of resources. The best-fit approach in one context may not always translate successfully in another. In all cases, it will be critical to keep in mind the core policy question driving the analysis as well as the broad range of data ecosystems in which users will be seeking to apply these methods, ranging from stable settings rich in frequent baseline and auxiliary data to fragile and conflict-affected settings with very limited data inputs and high implementation constraints.

1.

Methods for Nowcasting Welfare – *With a Focus on Monetary Poverty*

Nowcasting and imputation methods leverage baseline data that contains a direct measure of welfare and more recent auxiliary data sources with which welfare is imputed. The baseline data provides the foundation of the analysis, containing variables with which welfare can be estimated (for example, from a household survey). Auxiliary data sources vary; some models make use of household microdata such as labor force, census, demographic and health, or specially collected household surveys; others rely on more economy-wide data such as current GDP or Final Consumption Expenditure numbers. Some methods also use big data sources such as geospatial or call detail record data. Still, almost all these methods need the baseline information to understand how these auxiliary variables relate to welfare or require baseline income distributions to make inferences about changes in welfare.

In the following, several different methods of estimating welfare and poverty are described in more detail, with specific guidance on advantages and disadvantages, and example use cases and links to further resources are provided. Annex 1 provides a summary of the different methods, which are discussed in this typology, including requirements for the method to accurately estimate welfare indicators and what limitations the method has.

Before we move on, it is important to note that all models described next rely on important assumptions that need to be assessed and possibly validated in each context. When feasible, it is recommended to run different options to compare results. Triangulation of findings with other external sources of information is also advisable. Finally, all methods have errors, and wherever possible, confidence intervals should be reported with the results.

1.1 Nowcasting Welfare Using Survey and Other Non-survey Covariates

Survey and non-survey-based imputation methods (covariate-based nowcasting) model the relationship between consumption or income and other covariates to nowcast poverty. Survey-to-survey imputation methods draw upon distributions of consumption (or income) variables and other covariates from a baseline survey to nowcast consumption (or income) levels using a recent auxiliary survey, which itself does not hold consumption variables. Non-survey-based imputation draws upon information from non-survey auxiliary data, such as remotely-sensed geospatial data. These variables can either be used to improve survey-based models or to independently form imputation models.

While imputation across space has received considerable attention, advancements in survey-based imputation of welfare over time are still recent. Hentschel et al. (1998) and Elbers, Lanjouw, and Lanjouw (2003) initiated a wave of research within and outside of the World Bank to adapt imputation methods to estimate monetary poverty for poverty mapping. These imputation models have been widely used to generate spatially disaggregated welfare information.^{2,3} More recent work is exploring ways to adapt these models to update welfare across time.

Most commonly, linear regression models are used to impute consumption and expenditure variables. Survey and non-survey-based imputation models also involve statistical approaches like hot-deck imputation and multiple imputation (MI), which aim to reduce nonresponse bias and improve the overall representativeness and quality of survey data.⁴ Some studies estimate a poor or

² For more information about imputation methods across space, see for instance [Corral et al. \(2022\)](#), [Stifel and Christiaensen \(2007\)](#), [Tarozzi \(2007\)](#), [Christiaensen \(2012\)](#), [Mathiassen \(2013\)](#), or the report “[More Than a Pretty Picture: Using Poverty Maps to Design Better Policies and Interventions.](#)”

³ Survey-to-survey imputation can also be useful for other applications beyond updating or mapping monetary poverty, such as ensuring comparability of consumption over time or imputing non-monetary welfare metrics. Even when surveys are available, changes in poverty lines or consumption modules can hinder comparisons of poverty over time. Survey-to-survey imputation has also been used to restore comparability in such circumstances (see for example, Lain, Schoch, and Vishwanath (2023), or Fernandez, Olivieri, and Wambile (2024)). Moreover, imputation methods are applicable not only for monetary poverty estimation, but also for other non-monetary welfare outcomes if relevant predictors are available.

⁴ These methods relate to a larger literature in statistics on missing data or multiple imputation (see, for example, Rubin 1987; Carpenter and Kenward 2013). Compared to the standard survey-to-survey imputation framework, MI methods typically employ Bayesian estimation techniques, which are more

non-poor status of a household with a probit or logit model directly (for example, Mathiassen 2013), which however comes with stronger and more restrictive set of parameter assumptions. Furthermore, a linear model is, in most cases, more efficient, and a continuous dependent variable allows for a larger set of information to be included in the model (Dang and Verme 2019).

While most survey-to-survey imputation applications use linear models, non-parametric machine learning models could also be used to model the relationship between covariates and welfare indicators in the baseline and impute in the auxiliary data. Most machine learning applications are classification problems, imputing poverty headcounts directly, rather than predicting income or welfare measures.⁵ A key difference between most machine learning approaches and traditional methods are that the former are non-parametric statistical methods, while the latter are parametric regression models.⁶ The key is to choose the model with the best predictive power and lowest overfit. There are multiple types of machine learning methods that can be deployed, and they vary greatly in terms of their complexity and interpretability. Some methods, such as logistic regression or decision trees are easily interpretable but oftentimes do not offer the most accurate predictions. More complex methods, such as random forests, gradient boosting or support vector machines, may increase accuracy but are not as easily interpreted. The most complex and difficult to apply methods are those that fall under the bracket of deep learning. For example, in the context of image classification, while Convolutional Neural Networks have proven to be the most effective, their application requires substantial computing power and methodological knowledge. There can also be potential measurement errors.⁷ Moreover, every

computationally intensive and complex (Dang et al. 2017). Hierarchical Bayesian models have the advantage of enabling the specification of prior distributions in model parameters, accounting for spatial correlation, and providing estimates of uncertainty in the predictions (Steele et al. 2017). The choice of imputation method depends on the nature of the data and the specific goals of the survey analysis.

⁵ In machine learning terminology, classification problems refer to categorical variables, such as “poor” or “non-poor”, while regression problems refer to continuous variables, such as income or welfare measures.

⁶ Most commonly, machine learning methods are used for the prediction of welfare with non-survey data. This is particularly the case for large unstructured datasets, such as remote sensing imagery. However, non-survey-based imputation can also be conducted with “traditional” linear regression models, which are more often employed for survey-to-survey imputation. Likewise, machine learning methods can also be applied in survey-to-survey imputation models. For both survey and non-survey-based imputation, the best option may be to test different models and choose the one with the best predictive power, while having low overfit.

⁷ Classification procedures do not always correctly assign items in a satellite image. Mismeasurement can be mitigated by including non-remotely-sensed data into models (Alix-Garcia and Millimet 2023). Partial registries of household-level poverty predictors in Malawi vastly improve the performance of

model-based estimation of poverty comes with a degree of uncertainty. Corral et al. (2022) discuss some of the considerations and challenges of using machine learning for poverty nowcasting.

Examples of survey-to-survey imputation over time combine infrequent household budget surveys with more frequent labor force or demographic and health (DHS) surveys. For instance, Dang, Lanjouw, and Serajuddin (2017) compute poverty rates leveraging more frequent labor force surveys and baseline household survey data from Jordan. Household-level demographic and labor status information are used to predict consumption variables within the household survey data. In a second step, the model is then applied to the same predictors in the labor force survey data to impute consumption rates in the more frequently collected labor force survey. Cuesta and Ibarra (2017) also use a labor force survey for imputation for Tunisia. Doudich et al. (2016) use similar methods for labor force surveys in Morocco, while Stifel and Christiaensen (2007) use survey-to-survey imputations on DHS data in Morocco. Edochie et al. (2022) use both survey-to-survey imputation and growth-to-poverty elasticity methods to estimate poverty rates in India, with both methods yielding similar results. Beltramo et al. (2020) impute poverty rates among refugees in Chad from a UNHCR survey and administrative data.

While the academic literature predicting welfare or poverty rates using non-survey-based data and machine learning techniques has grown steadily, applications imputing welfare over time remain scarce. Non-survey datasets can, for example, be remotely-sensed data, such as satellite imagery of buildings or infrastructure (Jean et al. 2016, Yeh et al. 2020), or digital trace data, such as call detail records from phone users (Blumenstock, Cadamuro, and On 2015). Yet, studies that explicitly investigate changes over time are rare and at the time of writing, the authors are not aware of any studies that have been successful in reliably capturing changes in monetary or asset-based welfare over time.⁸ For instance, Browne et al. (2021) find that nowcasting asset-based poverty and malnutrition using DHS survey data and various geospatial data sources is much less

machine learning models that combine survey and satellite imagery for the purpose of village-level targeting (Gualavisi and Newhouse, 2022). Hence, in many cases traditional survey data can improve the accuracy of geospatial data for poverty and welfare prediction.

⁸ Yeh et al. (2020) use deep learning models based on satellite imagery and survey-based wealth estimations, which they find to predict 50 percent of the variation of district-level changes in wealth over time. In general, non-survey-based imputation models are found to achieve higher predictive accuracy when imputing wealth indicators than consumption variables, meaning that changes over time might be difficult to capture (Barriga-Cabanillas et al. 2022, Merfeld and Newhouse, 2023). Gualasavi and Newhouse (2021) show that collecting partial registries of villages improves the predictive accuracy of imputation models using geospatial data.

accurate than cross-sectional estimations, especially at more disaggregated levels of observation. Marty and Duhaut (2024) compare different models and data sources to predict poverty, finding models explain only between 4-6 percent of variation in asset-wealth over time (26 percent being the maximum in one country). Predicting non-monetary welfare indicators such as food security over time is more feasible than predicting monetary or asset-based welfare indicators (see for example, Andree et al. (2020) or Tang et al. (2021)).⁹

Blumenstock, Cadamuro, and On (2015) provide an application of non-survey-based imputation to update poverty estimates over both space and time by collecting additional auxiliary data. The authors predict district-level poverty rates in Rwanda using detailed phone record data. In absence of baseline data to link welfare measures with call detail record data, the authors collected survey data from a small sample of the phone customers, containing information on assets and other welfare indicators. With this data, Blumenstock, Cadamuro, and On (2015) trained a machine learning model to relate phone usage patterns to welfare indicators, and, in a second step, made out-of-sample welfare predictions for the whole sample of customers in the CDR data. This study shows how having access to real-time CDR data allows to nowcast poverty rates. While this particular application required the collection of new baseline data, in some instances, existing survey data can be linked to big data, for instance through household location (see for example, Jean et al. (2016)). Part 2 of this typology presents other examples in more detail.

Considerations Regarding Reliable Survey- and Non-survey-based Imputation

Model instability can occur if the relationship between consumption and the predictors in the model changes over time. The main assumption of survey-based imputation over time is that the joint distribution between consumption or expenditure and other covariates remains stable. Essentially, this assumption implies that any observed shifts in poverty levels over time are solely explained by changes in the model's included factors, rather than unobservable changes in the

⁹ Moreover, machine learning tools can significantly improve the accuracy of early warning systems (McBride et al. 2022). For instance, Lentz et al. (2019) apply LASSO techniques on remotely-sensed climate data, data on food prices, and demographic and asset data from LSMS surveys to make near-real-time predictions of food insecurity. The authors find that including spatially disaggregated demographic and asset data enhances model predictions. Browne et al. (2021) use health indicators and an asset index from DHS data, food price data, and climatic indicators to asset and malnutrition outcomes in near time using multivariate random forests.

returns of those factors (Dang, Lanjouw, and Serajuddin 2014; Corral forthcoming). Including variables that are good at predicting changes in welfare over time is important to improve the models, but does not ensure that the assumption of a stable relationship between covariates and welfare outcomes holds (Stifel and Christiaensen 2007; Newhouse et al. 2014). [Dang, Lanjouw, and Serajuddin et al. \(2017\)](#) developed a survey-to-survey imputation framework that has been validated in several different contexts, and that allows for distributions of estimated parameters to change over time. However, depending on the model used, differences between predicted and observed rates can be considerable (for instance, MI models performed worse than others in their study).

Another important condition is that predictors need to be comparable between the surveys. Different types of surveys (for example, household and labor force surveys) may have different definitions of variables used for imputation, such as household size or labor force participation. In such cases, variables do not follow the same distribution across surveys, violating a key assumption in imputation designs. For instance, Newhouse et al. (2014) and Dang (2021) find differences in estimated poverty from household and labor force surveys. Simple t-tests for equality of distributions can be used to test this condition (Dang et al. 2021). The Dang, Lanjouw, and Serajuddin (2017) framework includes procedures for the standardization of variables across surveys, should the condition be violated, which has been shown to be effective (Corral forthcoming). Basically, variables are rescaled by the relative differences of their variance and anchored to their means (see proposition 3 in Dang, Lanjouw, and Serajuddin (2017)). Harmonization across surveys can significantly reduce bias introduced by variables not being comparable.¹⁰ Yet, it is important to note that if differences between surveys come from different levels of representation or coverage biases, these corrections will not fully resolve the problem.

The selection of predictors is also key for good performance of these methods. Dang et al. (2023) examine the accuracy and robustness of survey-to-survey imputation using different types of predictors. First, imputation models that use household demographic and employment characteristics can be improved upon by including household utility consumption expenditure, assets, or dwelling attributes.¹¹ Resulting poverty estimates are statistically not distinguishable from true poverty rates. Second, geospatial indicators such as soil quality, distance to facilities, or

¹⁰ Another approach to harmonize variables across surveys is reweighting to match means (see for example, Roy and van der Weide (2022)). Reweighting may however still yield bias results, as it does not affect the distribution of variables (Corral, forthcoming).

¹¹ Including household utility expenditure increases the probability of obtaining accurate poverty estimates by 46 percentage points.

nightlights (in some contexts) further improve the accuracy of estimates. Third, the authors find that while these augmented models work well nationally and subnationally, there are some differences in model prediction accuracy between urban and rural areas. For urban areas, the inclusion of food, health, education, or utility expenditure items matters most, while in rural areas non-food expenditure or utility expenditure matters most. When there are two or more earlier rounds of data with consumption variables, Dang et al. (2017) propose an Oaxaca decomposition test to select predictors for the imputation model.

Survey-to-survey imputation methods have usually relied on data collected face to face. An area for further research is whether they can be used to leverage the mounting number of phone surveys. Household surveys administered via phone, which are discussed in more detail in part 2 of this typology, proliferated during the COVID-19 pandemic. Therefore, if recent baseline data exists and equivalent predictors are collected, survey-to-survey imputation methods could potentially fill this important gap. However, the very nature of phone surveys is that they are often collected when face-to-face surveys cannot be collected, for instance in fragile contexts where enumerators' access is limited or during pandemic lockdowns. Therefore, model stability assumptions are questionable and carefully need to be evaluated.¹² Boznik et al. (2017) found no reporting bias in predictor variables collected via phone, and that survey-to-survey imputed poverty rates are accurate in a study in Serbia. Phone surveys are shorter than standard household surveys, but reported consumption rates can be inaccurate in some circumstances (Abate et al. 2023). Yet, there is very little evidence on the performance of phone surveys for survey-to-survey imputation methods, an area which warrants further research.

Standard multiple imputation procedures can be biased if residuals do not follow a normal distribution. If the variable of interest does not follow a normal distribution in the baseline data, multiple imputation may lead to false predicted values in the auxiliary data. To achieve an approximate normal distribution, consumption, income or expenditure are typically log-transformed. However, in some cases a log-transformation may not solve the issue of a skewed distribution (Yoshida et al. 2022). Box-Cox transformation or log-shifts can be applied to recover normal distributions and improve imputation accuracy (Corral et al. 2021).¹³

¹² There can also be issues related to sampling bias and lack of representativeness, which are discussed in part 2.

¹³ For small area estimation, procedures such as a parametric estimation of the distribution of residuals have been developed to deal with non-normal distributions, which are, for instance, incorporated in the "sae" command in Stata (Corral et al. 2022).

Finally, models trained with data in one country are most often not transferable to other countries without loss of predictive power. The scalability of methods to predict welfare from mobile or other big data sources may be limited, as the accuracy depends largely on a specific country context. “Off the shelf” applications of machine learning algorithms trained on mobile phone data from one country to a different country are therefore not advisable in most cases (Blumenstock 2018).

Survey Imputation Methods can be Complemented with Data Collection to Deal with Missing Auxiliary Data

When the auxiliary data is not available or not usable due to risks of violating assumptions, a strategy can be to collect new auxiliary data. The Poverty and Equity GP’s Survey of Well-being via Instant and Frequent Tracking (SWIFT+) approach was developed for instances when there is no auxiliary data or when the comparability between the baseline and auxiliary data is limited. The key value-added of SWIFT+ is complementing survey-to-survey imputation methods with the collection of low-cost auxiliary data to estimate monetary poverty measures.¹⁴ The baseline data is used to identify the model and the main predictors of poverty, usually 10 to 15 variables, which are collected in a new survey for a representative sample. These variables include some additional fast changing variables for better modeling changes in poverty (such as food consumption items, employment variables, or economic sentiments) (Yoshimura et al. 2022). Because the auxiliary data collection is focused on a few variables, it comes at a comparatively low cost. The predictors in the short survey are framed in the exact same way as in the baseline survey, which minimizes errors due to lack of comparability between the surveys (Yoshida et al. 2022). After the covariates have been collected, a standard survey-to-survey imputation model (that is, linear regression models or machine learning models) can be applied to predict consumption or income.

Similar assumptions and considerations to the standard survey-to-survey imputation framework also hold for SWIFT+. Foremost, collecting new auxiliary data does not resolve issues with regard to the relationship between

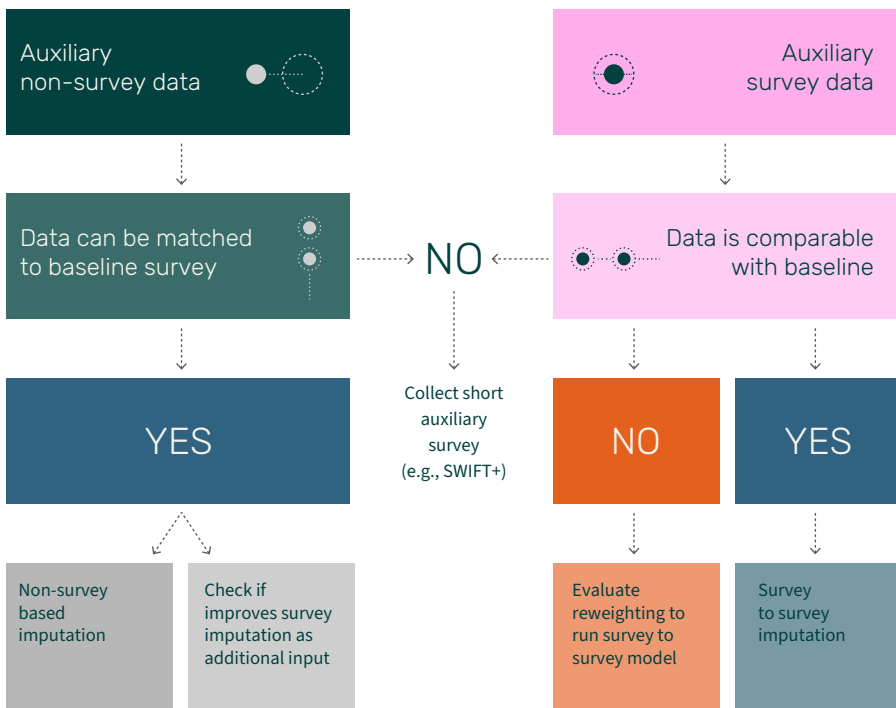
¹⁴ SWIFT+ is preferred to SWIFT 1.0 (the original SWIFT approach) as it allows for time-variant variables in the modeling. The original SWIFT 1.0 included only slow-changing variables, which reflect stocks better than flows. It was updated to incorporate fast-changing consumption variables, which is commonly referred to as SWIFT+ or SWIFT Plus. As capturing changes over time is of interest in most applications, and especially in the context of real-time monitoring of welfare, we do not elaborate on SWIFT 1.0 in this typology. For more details on SWIFT 1.0 see [Yoshida et al. \(2015\)](#).

covariates and consumption or expenditure potentially changing over time. As the auxiliary survey is collected and survey questions can be framed in a manner that is comparable to the baseline survey, concerns about variable comparability are however mitigated as long as the auxiliary data collection has the same coverage and representativeness of the imputation model. If the auxiliary data does not cover the same population as the model, an assumption must be made that the parameters of the model apply for the area/population covered by the auxiliary data.

Lessons Learnt and Resources

Survey-based imputation models are a powerful tool to produce updated welfare estimates. Figure 5 provides an overview of the main survey-based imputation methods discussed in this section in a simplified decision tree. When both baseline and auxiliary data exist, traditional survey-to-survey imputation methods can be used to produce more up to date estimates of welfare (if all assumptions are validated). When new auxiliary data is needed, SWIFT+ can be used.

FIGURE 5 Survey-based imputation methods simplified decision tree



Resources

Survey-imputation

- See [Dang et al. \(2017\)](#) for a readily applicable survey-to-survey imputation method that has been validated in several different contexts. The authors have made the *povimp* Stata package available, which can be used for imputations including different modelling options (Dang and Nguyen, 2014).
- For current reviews of survey-to-survey imputation studies, including guidance on modeling options and variables to include, see [Dang et al. \(2023\)](#) and [Dang et al. \(2019\)](#).
- See [Corral et al. \(2022\)](#) for guidelines on small area estimation for poverty mapping, and the *sae* package for implementation in Stata (Nguyen et al. 2018).
- [Yoshida and Aron \(2024\)](#) and [The Concept and Empirical Evidence of SWIFT Methodology](#) explain the methodology of the different SWIFT modules and applications in detail.
- The [SWIFT 1.0 Data Collection Guidelines](#) includes instructions for applications of SWIFT 1.0 in Stata.
- For more information on Rapid Consumption Surveys, [see chapter 9](#) in the handbook [Data Collection in Fragile States Innovations from Africa and Beyond](#).

Non-survey Imputation and Machine Learning

- See here a list of [World Bank blogs](#) on the topic of machine learning.
- See resources listed in section 2.2 on literature using non-survey imputation methods and geospatial data. Note that most of these are for small area estimation, rather than welfare prediction over time.
- Annex 2 includes various examples of studies that have made use of supervised machine learning methods and non-survey datasets to predict poverty rates in countries or regions without current survey data. Part 2 of this typology discusses caveats of machine learning models linked to [geospatial](#) and [CDR](#) data in more detail.

- The [Machine Learning for Disaster Risk Management](#) e-book by GFDRR provides a great introduction on machine learning methods and several examples of use cases.
- [Corral et al. \(2023\)](#) use real-world data to validate and compare different machine learning approaches for poverty mapping.
- Chapter 5 of the guidelines on small area estimation for poverty mapping by [Corral et al. \(2022\)](#) compare gradient boosting to traditional poverty mapping methods.
- See this [post and paper](#) on poverty mapping using machine learning by Stanford University researchers.
- The paper [Machine Learning Methods That Economists Should Know About](#) for an overview on the topic by Susan Athey and Guido W. Imbens (2019)
- On the topic of poverty prediction, see [A review of machine learning and satellite imagery for poverty prediction: Implications for development research and applications](#) by Hall et al. (2023).
- A discussion on [The Ethics of Machine Learning](#) in a blog post by Berk Özler.
- Another paper on conceptual approaches to Machine Learning: [Machine Learning: An Applied Econometric Approach](#) by Mullainathan and Spiess (2017).
- [Machine Learning trips and tricks cheatsheet](#) gives a short overview over different concepts, models and terminology by Afshine Amidi and Shervine Amidi (Stanford University).

1.2 Nowcasting Welfare Using GDP Growth

GDP-poverty nowcasts are a simplification of the complex dynamics of poverty and economic growth. The intuition behind these methods is simple and builds on the substantial evidence showing that across time and countries economic growth is strongly and significantly correlated with declining poverty rates, based on survey data (Ravallion 1995, Ravallion 2001, Dollar and Kraay 2002, Kraay 2006). There are two main approaches to GDP-based nowcasts: distribution scaling and the poverty-elasticity method.

The distribution-scaling method scales the welfare distribution as last observed by a fraction of GDP per capita growth. It is typically distribution neutral, scaling the whole welfare distribution by the same factor (that is, a linear growth incidence curve) (Kakwani 1993). Poverty rates are then derived from the updated levels of welfare. While economic growth can change the income distribution of a country (Ravallion 2001), most of the observed changes in poverty can be attributed to changes in mean incomes.¹⁵ World Bank estimates used for Global Poverty Monitoring typically assume distribution-neutral growth. Using distribution neutral changes in welfare since the last period has been found to outperform more complex models, such as estimating poverty directly based on current variables (Mahler et al. 2022a). Changes in inequality can nevertheless be integrated in the method by applying non-linear growth incidence curves, incorporating assumptions on how growth attributes to incomes of different households along the income distribution (Lakner et al. 2022).¹⁶

Shifts in the welfare distribution need to be adjusted by the discrepancy between income or consumption growth in household surveys and growth in national accounts.¹⁷ Research has found differences in income measured using national accounts data and household surveys (Ravallion 2003, Pinkovskiy and Sala-i-Martin 2016, Prydz et al. 2022). Therefore, a pass-through rate needs to be defined to translate GDP growth from national accounts into household income growth. Typically, a uniform pass-through rate is assumed. Mahler et al. (2022a) estimate separate pass-through rates for when welfare is measured from consumption or from income variables. Their estimates suggest that 71 percent of growth passes through to consumption aggregates and 97 percent passes through to income aggregates. Lakner et al. (2022) estimate a pass-through rate of 0.85 between real GDP per capita growth and welfare measured in 1429 survey spells. However, the pass-through rate is likely to be context and

¹⁵ Countries with higher inequality tend to exhibit a lower growth elasticity of poverty than countries with low inequality, meaning that economic growth is less beneficial for reducing the poverty rate in highly unequal societies (Ravallion and Chen 1997, Bourguignon 2003). Therefore, reducing inequality can directly reduce poverty today, if reductions in inequality benefit households below the poverty line, and accelerates poverty reduction from economic growth in the future (Bourguignon 2004; Alvaredo and Gasparini 2015). In fact, after 2010, the median elasticity of poverty with respect to inequality has been shown to be larger than the median elasticity of poverty with respect to growth (Bergstrom 2020).

¹⁶ Caruso et al. (2017) similarly propose a simple way of easing the assumption of distribution-neutral growth by applying different growth rates along the income distribution. These income-dependent growth rates are obtained by calculating the share that each income quantile contributes to overall growth between two periods in the past. This adjustment comes with the assumption that the contribution to overall growth of each quantile is stable over time.

¹⁷ In general, consumption data is preferred to income data, as it is more directly linked to economic welfare.

country specific. Lakner et al. (2022) propose a model-based recursive partitioning process, in which input variables (such as Gini coefficient, median income, or population) are interacted with real GDP per capita growth separately and Wald tests determine which variables have a statistically significant coefficient. The sample is split, based on the variable with the lowest p-value. The algorithm is then repeated by applying it to observations in each of the two subsamples separately. The result is a set of pass-through rates, that differ between countries depending on the identified input variables.

The poverty-elasticity method directly calculates how much growth in GDP per capita changes the poverty rate. This method computes the elasticity between poverty and GDP in the past and predicts current (or future) poverty based on changes in GDP.¹⁸ While it is possible to calculate elasticities based on only on two points in time, estimating elasticities based on multiple periods of GDP growth and poverty is favorable. Implicitly, the method is not distribution-neutral if in the past changes in GDP affected poverty, also through changes in the welfare distribution. To account for the fact that the elasticity between GDP and poverty will change over time depending on poverty levels, modeling elasticities should incorporate initial poverty levels (Corral et al. 2020).^{19,20} If elasticities are to be shown in terms of consumption or income instead of GDP, pass-through rates as explained above can be applied.

When reliable data on GDP per capita is missing, some studies have resorted to alternative data sources to impute changes in GDP. For instance, several applications have used nighttime lights data when GDP data is not available (section 2.2 will discuss nighttime lights data in more detail). Nighttime lights have been shown to be highly correlated with GDP (among other variables) (Chen and Nordhaus 2011, Henderson et al. 2012, Bruederle and Hodler 2018, Michalopoulos and Pappaianou 2018), which has proven useful when GDP values are not reliable (Martinez 2022).²¹

¹⁸ With more than two periods, regression models, ideally controlling for initial poverty levels, are used.

¹⁹ The authors show that GDP growth translates less to poverty reduction in fragile and conflict affected countries, than in other economies.

²⁰ A variant of the poverty-elasticity method is to use semi-elasticities instead, which refers to percentage point instead of percent changes in poverty, as a result of a percent change in consumption (Klasen and Misselhorn 2006). Effectively, this induces changes at lower incomes to be smaller, but changes at higher incomes to be larger.

²¹ Similarly, Fezzi and Fanghella (2021) use high-frequency electricity market data to estimate GDP loss during the COVID-19 pandemic in European countries.

Considerations Regarding GDP-based Nowcasting

Even though data requirements are smaller compared to other methods, not all countries have the necessary data. To calculate the elasticity between growth and poverty, at least two comparable poverty estimates from the past need to be available, which may not always be the case. Both methods require a relatively recent welfare distribution. Furthermore, not all countries publish national accounts. For instance, of all country-year observations without poverty estimates in the Poverty and Inequality Platform (PIP), only 53 percent have GDP per capita data available, and GDP numbers may not always be reliable in lower-income countries (Angrist 2022).

Estimates are subject to uncertainty. When only two points in time are available, extrapolating poverty elasticities carries high uncertainty and is not recommended. Second, an estimate that is too outdated may cause concerns that elasticities may have changed. When a recent welfare distribution is available, the scaling method is generally preferred (Caruso et al. 2017). Estimating poverty directly, as is done with the elasticity method, is outperformed by making use of the whole welfare distribution (Mahler et al. 2022a). Model instability can also arise when estimates from one country are transferred to another country. This can be the case for growth-poverty elasticities or pass-through rates, which are often estimated for a subset of countries and then applied globally. Corral et al. (2020), however, show that in fragile and conflict affected countries pass-through rates are much lower than in other countries.

The accuracy of both methods depends on the stability of the relationship between growth and poverty. Economic volatility or significant policy changes can change how growth affects poverty. Second, the poverty nowcasts based on GDP depend on a good estimation of GDP. A low-quality GDP projection can impact the reliability of the methods for estimating poverty. Importantly, as GDP numbers are published with a considerable time lag (Bańbura et al. 2013), GDP itself is typically imputed from household consumption or industry output data, which is then fed into poverty predictions.²² Moreover, changes in GDP and changes in poverty do not always move in expected directions. In about 30 percent of cases of all available data in PIP, GDP increases (declines) coincide with poverty rate increases (declines).

²² For more technical and methodological information on nowcasting, see for example, Giannone et al. (2008).

One of the largest issues with GDP-based nowcasting is that it does not capture other complex interactions very well. For instance, economic growth in unskilled labor-intensive sectors (agriculture, construction, and manufacturing) benefits poverty alleviation the most (Loayza and Raddatz 2010). Mahler et al. (2022b) use a variety of data sources to incorporate (some) distributional shifts in the GDP growth method, for instance, by assigning sectoral growth rates to different populations. For the poverty-elasticity method, this entails that if economic growth is highly sector specific, estimations can be improved by taking sectoral shares and growth rates into account (Mahler et al. 2022b). Ravallion and Chen (2007) and Loayza and Raddatz (2010) decompose economic growth by interacting each sector's value-added share in GDP with the growth in sector value added.²³ As employment status is highly correlated with being in poverty, a further extension of the elasticity method is to include employment elasticities in poverty estimations (see for example, World Development Report 2013).²⁴

Despite its limitations, nowcasting based on GDP has been shown to provide a simple and comparable way to update and project poverty across many countries. Recent research shows that nowcasting poverty rates based on GDP and welfare distributions outperforms alternative more complex models, even when the last survey is up to five years old (Mahler et al. 2022a).

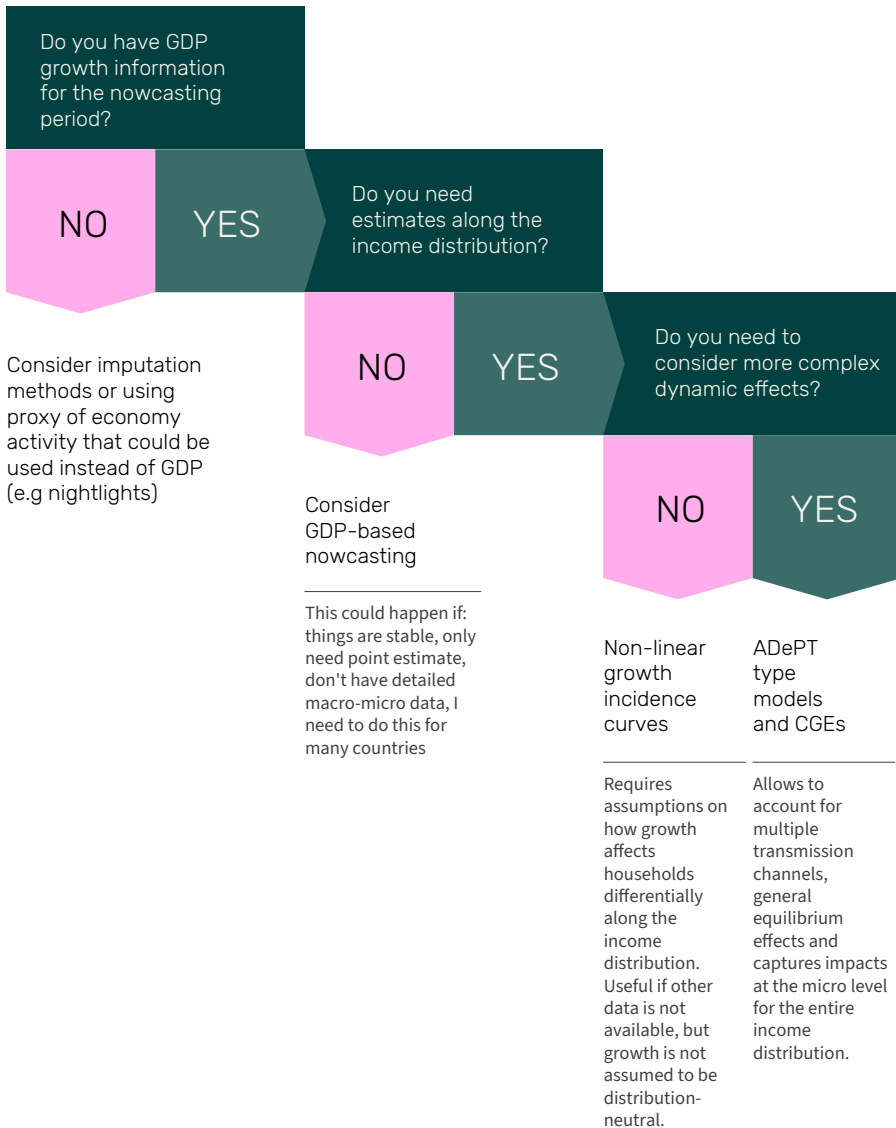
Figure 6 presents a simple diagram that can help navigate the different options in this and the following section. If the goal is to produce a point estimate of poverty and GDP data is available, GDP-based nowcasting models can be considered. This may be a particularly good fit when there is a need to provide a comparable way to update and project poverty across many countries or when more detailed macro data is just not available. While the relative simplicity of these approaches allows for greater cross-country comparability, it also comes at a cost. If there is a strong need or interest in considering broader distributional changes, non-linear growth incidence curves can be applied. However, if no reasonable assumption can be made on the shape of growth incidence

²³ Relatedly, PovStat—a Microsoft Excel® based program from the World Bank—simulates the poverty implications of alternative growth paths of specific economic sectors. The model does not incorporate other distributional effects beyond broad economic sectors, and only accounts for employment of the household head and neglects non-labor income. Therefore, poverty at the household or individual level cannot be computed (Habib et al. 2010). Changes in household welfare are assumed to evolve according to changes in per capita output in the sector of employment of the head of the household (Essama-Nssah 2005).

²⁴ Viollaz et al. (2023) however find that using employment elasticities underestimates income declines after a shock as compared to using actual employment data.

curves, or the goal is to account for dynamic effects, researchers may need to rely on microsimulation models, which are discussed in more detail in the following section.

FIGURE 6 GDP-based nowcasting and microsimulations simplified decision tree



Resources

- Data on income distributions is available on the [Poverty and Inequality Platform](#). The *povsim* Stata package can be used for poverty nowcasting, including linear and non-linear incidence curves (Lakner et al. 2014; Lakner et al. 2022).
- The World Bank’s biannual [Macro-Poverty Outlook](#) series and the [Poverty and Shared Prosperity](#) Reports use nowcasting and forecasting methods to estimate poverty rates globally, based on macroeconomic indicators.
- See [Yoshida et al. \(2014\)](#) and [Caruso et al. \(2017\)](#) for a discussion of both GDP-based methods.
- [Mahler et al. \(2022a\)](#) and [Mahler et al. \(2022b\)](#) describe the distribution-scaling method, which is also used in the Poverty and Shared Prosperity Report 2022.
- [Corral et al. \(2020\)](#) use and explain the distribution-scaling and the growth to poverty elasticity methods.
- [The Poverty and Distributional Impact of Macroeconomic Shocks and Policies: A Review of Modeling Approaches](#) includes an overview and discussion of several models for distributional analysis.

1.3 Nowcasting Welfare Using Microsimulations and General Equilibrium Models

By drawing on richer information that cannot be captured well by methods relying only on changes in GDP, microsimulation models can provide a more accurate estimation of the poverty rate (Mahler et al. 2022a). Microsimulation models draw upon rich data to disentangle how macroeconomic changes affect different households along the income distribution. Such granular data make it possible to account for different transmission channels and distributional effects, which can significantly improve poverty estimates (Montoya Munoz et al. 2023). These methods may provide more accurate nowcasts when the relationship between changes in GDP and welfare is not the same across the income distribution, when there are specific types of shocks that have heterogeneous impacts on different households, and when there is reason to believe that the relationship between baseline data and auxiliary data has changed over time.

Compared to the GDP-based models, some of these models are considered bottom-up, as they simulate actions and behavior of agents in the economy. Models that project macroeconomic changes onto microeconomic distributions are called top-down macro-micro simulations. Linking aggregate variables (LAVs), such as changes in prices, wages, and employment, form the link between the macro- and micro-models.

There is a range of macro-micro models that differ in complexity and coverage. Some macro-micro and simulation models aim to capture economic interlinkages as closely as possible to account for various direct and indirect effects that shape welfare outcomes of households and individuals, incorporating country or regional-level computable general equilibrium (CGE) models.²⁵ In such cases, macroeconomic shocks feed into general equilibrium models, which are linked to micro-level distributional inputs. Other models are more toned down in the light of the reality of the data landscape in many countries, but can still incorporate welfare impacts along the income distribution, expanding simpler growth-to-elasticity models. Instead of complex CGE models, simple macroeconomic projections are taken from macroeconomic forecasts, which form the macroeconomic impact in the simulation model. In the following, we summarize the most commonly used simulation models focused on consumption or income distributions and poverty in the World Bank.

The ADePT model simulates the distributional impacts of a macroeconomic shock, with relatively limited data and computational requirements. The [ADePT simulation model](#) is a simple macro-micro model, which projects how macroeconomic changes affect households through a variety LAVs, such as labor and non-labor income or food and non-food prices. It is particularly useful when no CGE models exist for the country of interest or if there is no interest in modeling dynamic effects. It can incorporate behavioral responses of individuals and households, such as adjustments in savings or occupation (Olivieri et al. 2014).²⁶ Figure 7 illustrates the modeling process of ADePT (Olivieri et al. 2014).

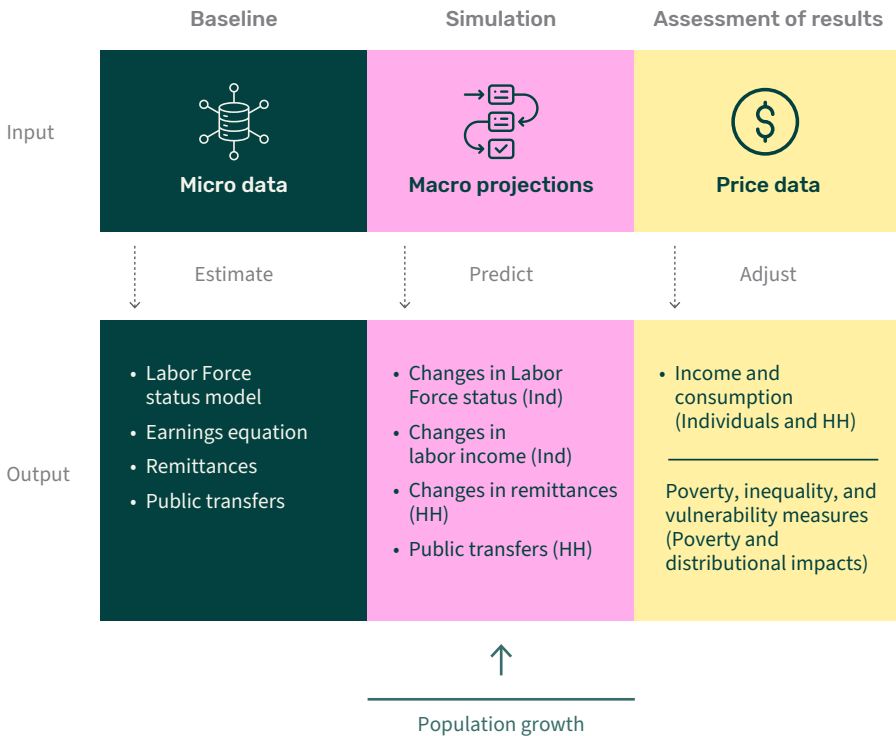
There are national and regional extensions to ADePT, which for instance allow modeling exercises for several countries at once or that incorporate more complex general equilibrium models. For example, in the Latin America and the

²⁵ CGEs simulate the market economy, including labor, capital, and commodity markets, and provide a tool to measure how specific changes transmit to the economy through prices, activities, and factors of production. In the general equilibrium, models capture behavior of households, firms, and the government in response to these shocks.

²⁶ Other types of macro-micro models involve iterative convergence methods based on such behavioral responses, which do not impose a top-down relationship but are much more complex to solve.

Caribbean region, the Poverty and Equity GP has set up simulation tools to project changes in labor market structures and incomes, which includes CGE modeling (Montoya Munoz et al. 2023).²⁷ Similarly, for the Africa West region, a model is being developed that can be run for 10 countries based on harmonized data. The model separates changes in household nominal incomes and cost of living, allows for heterogeneity in sector of income, and estimates exposure to inflation based on consumption shares. Estimated changes in household welfare come through projected changes in sectoral GDP and inflation rates. Other extensions include incorporating internal migration in Iraq, and non-constant income-consumption shares in SAR.

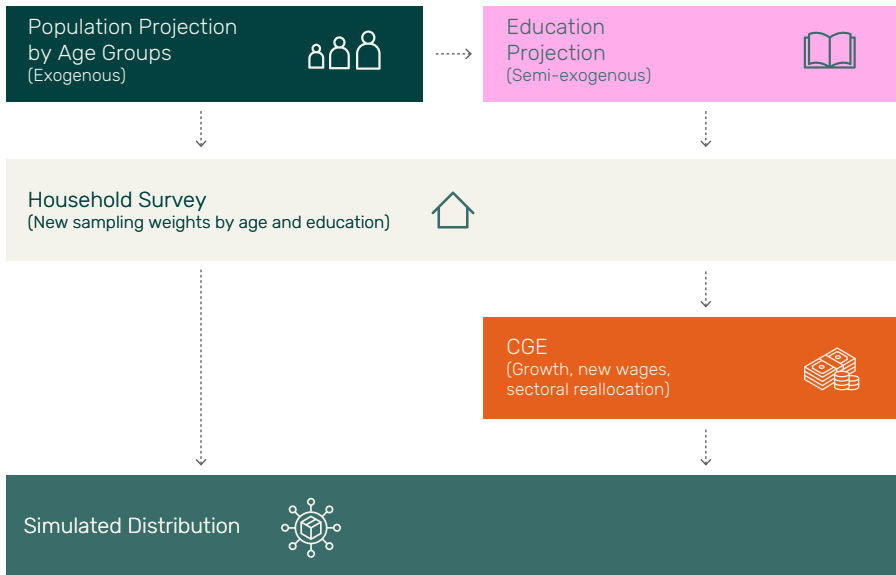
FIGURE 7 ADePT microsimulation process, World Bank LAC



Source: Olivieri et al. (2014).

²⁷ The model includes job losses, labor income changes, and non-labor (remittances) income changes as LAVs and has recently been updated to include intra-sectoral variation in wages depending on formality status, which improves model predictions in the medium term (Montoya Munoz et al. 2023).

FIGURE 8 The GIDD methodology



Source: Bourguignon & Bussolo (2013).

The Global Income Distribution Dynamic model (GIDD) is a global simulation model that ensures cross-country comparability and enables addressing a broader set of research questions. The GIDD framework synthesizes household survey data from more than 120 countries (Bourguignon and Bussolo 2013). Like ADePT, it calculates baseline income distributions, which are then linked to a global CGE model. The GIDD model can follow a dynamic process, modeling how baseline household characteristics, such as education or household composition, change over time in line with demographic projections. However, to estimate short-run effects, which are more relevant for real-time monitoring purposes, static assumptions are usually sufficient. Figure 8 provides an overview of the functioning of the GIDD-CGE model. The CGE model used in the GIDD process is the World Bank LINKAGE model. GIDD can also be used to assess the effects of global trade integration or of climate change.

Considerations Regarding Microsimulation and General Equilibrium Models

In general, microsimulation models are based on a set of assumptions, which may be overly strict in specific circumstances. For instance, in most models, labor and capital is modeled to be immobile, while shocks may cause workers

to move from one place to another. The models further assume that the relationships between variables remain stable over time and are not affected by the shock themselves. Coverage and nonresponse bias in baseline surveys can have severe implications (Figari et al. 2015).²⁸ Furthermore, microsimulation and general equilibrium models require setting multiple parameters, which affects modeling outcomes. As discussed for GDP-based nowcasting, macroeconomic data is not always reliable.

Assessing the accuracy of complex microsimulation and general equilibrium models is rare and difficult. Evaluating the predictive power of CGEs out of sample is close to impossible because they often rely on many context-specific assumptions and require vast amounts of data that may not be available in all settings. Moreover, as simulation and general equilibrium models include many channels that can be affected by many factors, such as policy changes, technological advancements, or other external shocks, out-of-sample validation is challenging. Judging the predictive power of such models is therefore difficult.

A common use case for real-time monitoring is to assess the impact of shocks, such as natural disasters, on poverty and welfare Annex 4, summarizes a discussion on how vulnerability and damage function models relate to real-time monitoring of welfare.

Resources

- The book [The Impact of Macro Economic Policies on Poverty and Income Distribution](#) presents the foundations of macro-micro evaluation techniques and tools, including top-down approaches and general equilibrium models.
- The report [The Gradual Rise and Rapid Decline of the Middle Class](#) in Latin America and the Caribbean has several applications of macro-microsimulation models. Annex 5 in the report explains the inputs and methodology of these models.
- The World Bank's biannual [Macro-Poverty Outlook](#) series uses microsimulation models to nowcast country-level poverty rates, if data for the country is available.

²⁸ Non-take-up and non-compliance are also factors that distort results.

- The methodological section “Introduction to Use of the ADePT Simulation Module” by Olivieri et al. (2014) in [Simulating Distributional Impacts of Macrodynamics: Theory and Practical Applications](#) provides an overview over the ADePT model.
- For more information on [the CGE-GIDD model visit the Poverty and Equity intranet website](#).
- For more information and Stata code for [the ADePT simulation model visit the Poverty and Equity intranet website](#).
- See this page on [global tools for Country Climate Diagnostics](#), which includes various climate simulation models.
- [The Poverty and Distributional Impact of Macroeconomic Shocks and Policies: A Review of Modeling Approaches](#) includes an overview and discussion of several models for distributional analysis.
- See [SOUTHMOD](#), a tax microsimulation tool, which is used to compare impacts of different policies on poverty and inequality. Another useful tool is The [Household Impacts of Tariffs \(HIT\)](#) simulation tool, with which users can simulate how changes in import tariffs impact the incomes of households across the income distribution.

2.

Harnessing Data for Real-time Welfare Monitoring

Timely data is a critical input to produce updated information on welfare. The rise of digitalization and new technologies have expanded the types of timely data that are feasible to collect or available to draw from. In terms of collecting data, greater access to mobile phones and internet connectivity has made it possible to use remote data collection modes (for example, phone, online, SMS) in many developing country settings. In terms of new data that is available to draw from, big data sources such as geospatial data, digital trace data, and administrative data continually provide updated information on how different variables are evolving. These sources, while not a replacement for traditional face-to-face surveys, can provide information to enhance the value of existing survey data or to reduce the sole reliance on that survey data.

There are three ways in which timely data is useful for real-time welfare monitoring. First, timely data provides inputs into models discussed in section 1 of this report, which can fill in missing data or complement existing survey data to nowcast welfare. Second, even when data is not sufficient to produce poverty estimates, timely data can provide up-to-date information on different welfare dimensions or proxies of welfare. For instance, changes in food security, employment status or subjective well-being are all relevant measures of (non-monetary) welfare. These variables are easier to collect, as they do not require a detailed household budget survey. Third, leading indicators can signal changes in welfare before these are detected by survey data. Examples include predictions on drought or floods, inflation forecasts, or using text-mining methods on newspaper articles to predict food insecurity (Balashankar, Subramanian, and Fraiberger 2023). Figure 9 provides an overview of the types of real-time data that are discussed in this section, and their potential use cases.

FIGURE 9 Data for real-time welfare monitoring—types and use cases

| Types of data | Inputs in models as auxiliary data | Proxies of welfare measures | Leading indicators |
|---------------------|------------------------------------|-----------------------------|--------------------|
| Phone survey | ✓ | ✓ | |
| Face-to-face survey | ✓ | ✓ | |
| Online survey | | ✓ | ✓ |
| Geospatial data | ✓ | ✓ | ✓ |
| Digital trace data | ✓ | | |
| Administrative data | | ✓ | |

Note: ✓ indicates that there are fairly successful applications in the area. The yellow shading for using phone surveys as inputs into auxiliary models reflects that there have been successful applications but more evaluation is needed.

2.1 Rapid Survey Data Collection

High-frequency Phone Surveys

Main characteristics and examples

High-frequency phone surveys (HFPS) have become popular, especially in contexts where in-person data collection is difficult. Phone surveys are a useful tool for monitoring purposes as they can be collected more quickly, at a much higher frequency and lower cost than large-scale face-to-face surveys. In Africa, the cost of HFPS was 30-times lower than a face-to-face survey (Zeza et al. 2022). However, their scope is usually much reduced, and they do not include extensive consumption and expenditure modules. As with other surveys, they can be collected either as cross-sections or as short or medium-run panels (re-interviewing the same individuals or households over time).

While HFPS have been used for some time, the COVID-19 pandemic expanded their use dramatically. The pandemic health risk and mobility restrictions led to a need to collect information by phone. For example, the Poverty and Equity GP conducted more than 500 phone survey rounds in at least 90 countries across all regions,²⁹ which can be accessed through the [COVID-19 Household Monitoring Dashboard](#). HFPS included different modules and areas of focus related to health impacts, vaccinations, economic impacts, education impacts, among others. These surveys conveyed important policy-relevant information; for example, in Ethiopia they showed that the number of people living in extreme poverty in urban areas increased by 33 percent (see Figure 10 below). HFPS have further been used to collect information on the impact of the pandemic on businesses, such as with the World Bank [Business Pulse Surveys](#). After the COVID-19 pandemic, phone surveys have also been used to track the implications of high inflation or other events across countries.

High-frequency phone surveys have been implemented to track various indicators related to welfare and perceptions of well-being. For example, the series of “Listening 2” surveys were designed to provide updated information on social and economic well-being every month to support welfare monitoring as circumstances change (for example, in the wake of crises or policy reforms). The results were able to be reported with a very short time lag, in some cases even within a week. These surveys have been successfully implemented in the Europe and Central Asia (ECA) region.³⁰ The [Listening 2 Tajikistan](#) perception surveys have been running since 2015, and similar efforts have been implemented in [Uzbekistan](#), [Kazakhstan](#), and [Kyrgyzstan](#). A key feature of the ECA surveys is that they are a representative panel, tracking information from the same respondents over time with overall low attrition rates. Such panels are particularly useful in tracking changes in welfare, economic status, and well-being over time. When a shock happens in between survey periods, its effects on households and individuals can directly be estimated. Core modules include public services, shocks, socio-economic well-being, jobs and employment, incomes, migration, and views on policy reform and development priorities. With the ongoing war in the Ukraine, Listening 2 phone surveys are being implemented to monitor people’s well-being and how the war impacts their livelihoods. These efforts demonstrate the usefulness of collecting proxy indicators to inform governments and international stakeholders.

²⁹ See for instance the work in [Latin America](#), [Sub-Saharan Africa](#), and [South Asia](#).

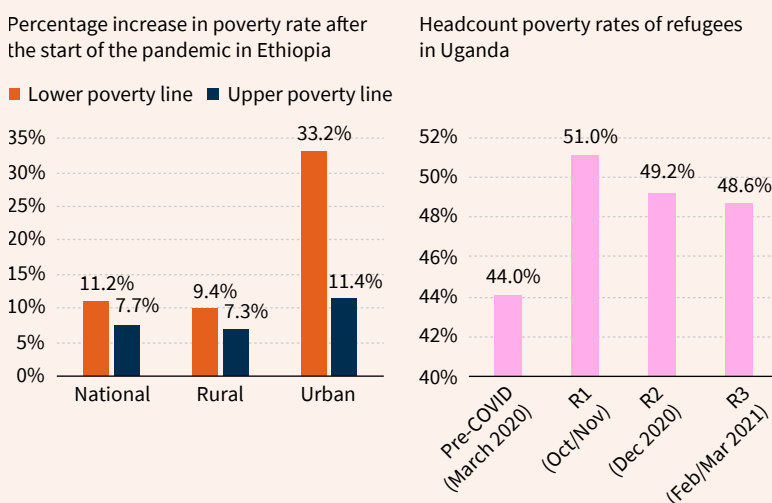
³⁰ Earlier efforts—not as successful due to high non-response—were piloted in [Africa](#) and Latin America.

Box 3 Examples of phone surveys in fragile and conflict-affected regions

Data collection in fragile and conflict-affected regions is difficult. When the safety of enumerators cannot be ensured, alternatives to standard face-to-face surveys need to be explored. Phone surveys have proven useful in such contexts. Examples of HFPS include tracking displaced populations in Mali and monitoring changes in household welfare and the situation of women in [Afghanistan](#). They also have been useful in monitoring the situation of hard-to-track or mobile populations, such as refugees or internally displaced persons, who carry their phones with them (Kim and Tanner 2023). [This](#) blog post summarizes some lessons learned from using HFPS to monitor the situation of displaced populations.

For instance, HFPSs in [DRC](#) (13 rounds), Uganda, [Somalia](#), and [Burkina Faso](#) (three rounds each) tracked internally displaced, refugee, and vulnerable populations. These surveys made it possible to monitor the status of these populations, and their vulnerability to further shocks. Findings suggested that the poverty rate among refugees was high, and the situation was exacerbated during the COVID-19 pandemic (see Figure 10).

FIGURE 10 Data collection in vulnerable settings



Source: [COVID-19 Household Monitoring Dashboard, Monitoring Social and Economic Impacts of COVID-19 on Refugees in Uganda: Results from the High-Frequency Phone Survey – Third Round.](#)

Caveats for high-frequency phone surveys

Despite their increasing popularity and appeal for monitoring, HFPS are not always a good substitute for representative face-to-face surveys. Especially in the context of collecting data on poverty and welfare, some disadvantages of using HFPSs arise. A few considerations are summarized next.

First, phone surveys need to be significantly shorter, simpler, and focused, compared to face-to-face surveys. Researchers need to decide what key areas they want to understand and design an instrument that does not exceed more than 15-20 minutes. Moreover, questions need to be simple so the respondent can answer them by phone. Respondents are more likely to quit interviews or lose concentration when interviews become too lengthy, especially in phone surveys (Abay 2021). A study in Ethiopia showed that poverty rates collected by a phone administered consumption survey are much lower than in a face-to-face survey, and that survey fatigue started much earlier in the former than the latter (Abate et al. 2023).³¹ This makes collecting accurate poverty data, which is typically based on a full set of consumption and expenditure variables, difficult. Using survey-based imputation (for example, with SWIFT+ modules) can partly address this caveat, when a full baseline survey exists.

Second, the surveys cover people who own a phone and have a connection. Even though more and more people around the world have phones, still only two-thirds of the population in low-income countries has a mobile phone (ITU 2023). Among those who do have phones, the affordability of using mobile services may be a constraint for low-income populations. Hence, surveying only populations with a phone may result in coverage bias towards wealthier, urban, and more educated households (Himelein et al. 2020, Ambel et al. 2021). Kugler et al. (2021) find that household heads, male, and older respondents are most often overrepresented in phone surveys in South Asia, Middle East and North Africa and East Asia and Pacific.

Third, conditional on focusing on a phone-owning population, the applied sampling method determines the extent to which the surveys are representative. The first best sampling strategy is to base the phone-survey sample on an existing representative (face-to-face) survey, such as the LSMS or DHS surveys (Himelein et al. 2020). This was implemented in many COVID-19 surveys (for example,

³¹ Not only respondents, but also enumerators experience fatigue in longer surveys, increasing measurement error (Finn and Ranchhod 2017).

Nigeria, Bangladesh urban and Cox’s Bazar surveys, etc.). In the absence of an existing frame, an extensively used strategy was to rely on Random Digit Dialing (RDD) (that is, to call a random list of phone numbers) (Brunckhorst et al 2023a). The reliability of this approach varies significantly depending on how the survey is implemented (a good practice example is the [SAR regional COVID-19 survey](#), which implemented several quality control measures and strategies for obtaining balanced samples across population groups). The third option is to randomly select numbers from telephone lists of telecom providers, which in general is faster than option 2, but comes with the same caveats and necessitates establishing arrangements with telecom companies.

Fourth, response rates are lower in phone compared to face-to-face surveys. Phone numbers are disconnected, electricity or network coverage may be faulty, or respondents cannot be reached by phone or incoming calls are not trusted, leading to nonresponse bias (Himelein et al. 2020, Brubaker et al. 2021). Incentivizing participation can reduce nonresponse bias, though highly frequent surveys may lead to response fatigue. In the COVID-19 HFPS, attrition rates were, however, relatively low (Gourlay et al. 2021). There are moreover significant differences in response rates between RDD and survey-based sampling. For instance, a World Bank study found response rates in survey-based sampling were 73 percent compared to 36 percent in RDD surveys (Brunckhorst et al. 2023b).³² It is also important to pay attention to language barriers. The SAR regional COVID-19 survey provides a useful reference on this area as summarized in this [blog](#).

A further bias of phone surveys is that they are skewed towards household heads and in some cases a particular gender. When information on the individual rather than the household level is of interest, phone surveys may not be suitable (if there is only a limited number of phone connections within the household), as household heads are much more likely to answer phone surveys than other household members (Brubaker et al. 2021). As household heads are more often male, there often is also a gender imbalance in survey respondents (not controlling for household head status).³³ Household heads for instance are more likely to be wage employed and better educated than the overall population (Ambel et al. 2021, Kugler et al. 2021). This drawback can be mitigated when information,

³² In panel studies, attrition rates are also found to be higher in RDD as compared to survey-based samples.

³³ In COVID-19 HFPS in eight South Asian countries, half of respondents were asked to pass the phone to an adult female household member to increase the share of female respondents.

such as labor market status, is collected for all household members, as was done, for instance, in HFPS in Nigeria and Madagascar, or by asking to speak to other household members explicitly. In some countries, gendered social norms may constrain who is able to answer the phone, which also creates a source of bias. However, similar concerns also arise in face-to-face surveys.

Different weighting adjustments can be used to improve the representativeness of the phone surveys. When the phone survey sample is based on an existing representative (face-to-face) survey, the survey data can be used for sampling weight adjustments. Demographic and socioeconomic characteristics of both of respondents and non-respondents of the phone survey can be used to adjust sample weights (Gourlay et al. 2021; Lee 2006, Kugler et al. 2023, Zhang 2023). There are two main methods in adjusting weights. The first, **weighting class adjustments**, splits the phone survey sample into cells by observable characteristics that could be correlated with the response-status (such as gender, education, location, age, etc.). For each of these cells, the survey weight of interviewed households is multiplied with the inverse phone survey response rate (Little 1986, Ambel et al. 2021). One advantage of this type of adjustment is that it does not necessarily require micro-data but can also be implemented if broader summary statistics of the population are available. The second method is **propensity score adjustments**, which is a model-based technique.³⁴ The actual response status of everyone is regressed on observable characteristics (using a probit, logit, or linear probability model). The inverse of the resulting predicted probabilities is then multiplied with the survey weight (Morgan and Todd 2008, Schafer and Kang 2009, and Austin 2011). These individual and more aggregate level weight adjustments can also be combined or sequenced (Zhang 2023).³⁵

While weight adjustments can reduce sample biases when information about the representative underlying sample is available, they still do not always make surveys representative. Extensive literature on this shows that non-sampling bias is not only determined by observed characteristics. When surveys are drawn from RDD or telecom provider lists, data from the sampling population is typically not available, which makes the re-weighting even more challenging.

³⁴ Most COVID-19 HFPS use propensity score reweighting techniques to re-calculate survey weights.

³⁵ There are also other methods of sampling adjustments, such as calibration of nonresponse (see Himelein et al. (2020) or Ambel et al. (2021) for further details).

Rapid face-to-face Surveys

In some settings, it is possible to conduct face-to-face surveys in a more timely and frequent manner by narrowing their scope and streamlining data collection. This comes at the cost of not having full survey information on all households and having to rely on imputation methods. Another way of reducing collection time and cost is to leverage local resources and streamline the survey implementation process. Rapid in-person surveys can be also useful to assess important specific focused questions such as household dynamics in the face of natural hazards (Knippenberg, Jensen, and Costas 2019).

Within-survey imputation as a strategy to shorten face-to-face surveys

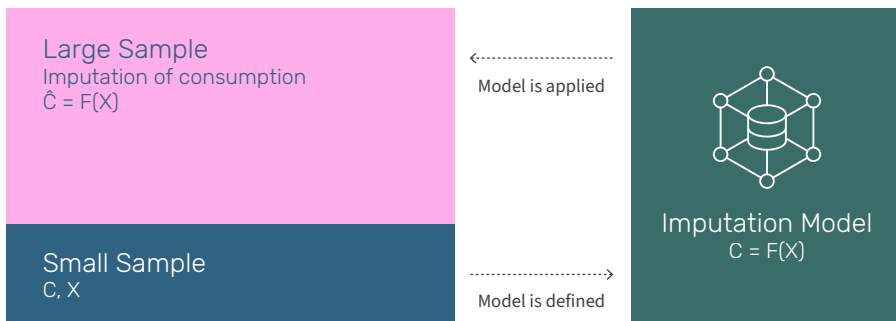
Within-survey imputation models can be used to reduce the length and associated cost of data collection. Within-survey imputation models rely on collected consumption from only a small part of the sample, which is used to train imputation models to predict consumption based on covariates collected for the rest of the sample. When there is recent auxiliary data with relevant variables, but there is either no baseline data or the baseline data is not comparable, only a small HBS needs to be collected, on which the model is trained that is used for imputation in the existing auxiliary data.

One example is SWIFT 2.0 which leverages within survey imputation models to newly collected baseline (and, sometimes, also auxiliary) data. SWIFT 2.0 trains a machine learning model on a (sub-)sample that contains the expenditure data. The model is then used to predict poverty for the larger survey that does not include expenditure data (Yoshida et al. 2022). This procedure reflects a tradeoff between collecting detailed consumption and expenditure data and reducing survey time and costs (Figure 11). The SWIFT 2.0 framework has been used to impute “real-time” estimates of monetary poverty in a variety of settings (see Yoshida and Aron (*forthcoming*) for examples).

The Rapid Consumption Survey (RCS) approach is another within-survey imputation method that aims to cut the cost of data collection for poverty measurement. The largest hindrance to implementing face-to-face surveys frequently and more swiftly is the large number and complexity of questions that are needed to estimate poverty rates. With RCS, enumerators administer a core module with the most important consumption items for all households. The remaining items are randomly allocated into separate modules, which are assigned to households. This reduces interview time by more than half, to

roughly 45 to 60 minutes per household. For each household, items from the missing module are imputed from other households, based on the core module and on household characteristics. Bootstrapping procedures are used for multiple imputation to correct for underestimation of household consumption (Pape and Mistiaen 2018; Pape 2021).

FIGURE 11 SWIFT 2.0 within-survey imputation



C: Consumption variables (collected) | **X:** Consumption correlates (collected) | **C-hat:** Consumption (imputed)

Compared to SWIFT 2.0, RCS collects a smaller set of consumption data for the whole sample, instead of a full consumption survey just for a subsample. Whether SWIFT 2.0, a full survey or RCS is more cost-effective depends on the overall sample size. When the overall survey sample size is small (below 600 individuals), RCS is more cost-effective than SWIFT (Yoshida et al. 2022), while the cost advantage of SWIFT 2.0 increases with sample size. Note also that RCS cannot be added to an existing survey, while the SWIFT-baseline module could be used to complement an existing auxiliary survey. Table 1 in part 2 of this typology reviews interview costs of different survey types.

Both approaches are superior to simply leaving out certain items completely, which can lead to the underestimation of poverty (Beegle 2012). Reducing the number of items in the core module and increasing the number of optional modules decreases survey time, though at the cost of accuracy. Furthermore, the selection of core items is important and ideally is based on an assessment of earlier survey data in the same country. When no data is available or when consumption patterns have changed markedly vis-à-vis the last survey, the performance of the estimator worsens (Pape 2021). Bias could also arise when the ordering or the mentioning of other consumption item matters for the response on an item.

Engaging local enumerators to lower face-to-face survey costs and increase the frequency of data

An example of rapid face-to-face data collection is the Malawi Rapid and Frequent Monitoring Systems (RFMS), where survey data is collected monthly by locally recruited enumerators. A collaboration between the WB Poverty and Equity GP, USAID, FCDO, Catholic Relief Services (CRS), Cornell University, and the Malawi National Statistics Office, implemented the RFMS to obtain real-time and frequent information on vulnerable populations. Enumerators were hired through a collaborative recruiting process in local communities and were typically youth who were active in community groups (Yoshimura et al. 2022). The approach proved to be more affordable (\$1.50 per household) than traditional and larger surveys (\$300-400 per household) and high-frequency phone surveys (\$10-40 per household). One large factor in cost reduction is that transportation and lodging expenses are lower than in traditional surveys (Taptué and Hoogeveen 2020). However, a governing body needs to be established to oversee data collection, which incurs a large fixed-cost. Therefore, community-based surveys have a cost advantage when surveys are collected at a larger scale and at a higher frequency. Naturally, survey and enumerator costs depend on the country and context of implementation.

The Malawi RFMS made it possible to track welfare indicators at high frequency. Among other findings, the data revealed seasonality patterns of hunger and showed how hazardous events magnified food insecurity.³⁶ An important advantage of routine data collection is that unexpected shocks can be monitored and vulnerable populations can be identified more quickly than when new surveys need to be collected. Furthermore, it allows for analyses over longer time horizons, as compared to being restricted to ex-post evaluations.

However, community-based surveys may be difficult to scale. The Malawi RFMS is collected in 10 districts of Southern Malawi. Community-based surveys are cost-effective when there are local enumerators, which might not be the case in all regions or urban areas. Furthermore, survey costs will be higher in higher-income countries and regions. It also needs to be considered that respondents may give different answers to some survey questions, if the enumerator is from their community and is someone that is known to them.

³⁶ Collected data is fed into a survey-to-survey imputation model (SWIFT+), allowing to estimate poverty rates. Having this type of ongoing data collection system in place made it possible to monitor the impacts of COVID-19 as well as the impacts of two tropical storms that occurred in Malawi. The storms exacerbated food insecurity and suppressed the typical reduction in poverty during the harvest season in the summer months.

Whether high-frequency data should be collected face-to-face or via phones depends on the specific context. In regions with lower mobile penetration rates, face-to-face surveys may be preferable, to counteract large sampling biases. Hiring local enumerators can also help to defer trust issues of respondents. On the other hand, finding suitable local enumerators to collect data repeatedly over longer time horizons can be challenging, especially in areas characterized by insecurity and vulnerability.

Online and Messaging-based Surveys

Web surveys are sometimes used to reach many respondents at low cost. Surveys are pre-programmed, and respondents fill them out independently. Web surveys can be sent via text or email to a specific target group or can be published online for anyone to complete (some websites offer small monetary remunerations for completing surveys). However, the issue of representativeness is aggravated with web surveys due to endogenous selection of people having access to the internet and taking part in web surveys, leading to severe coverage and non-response bias (Bethlehem 2010).

Social media platforms, such as Facebook, have been used as a platform to administer web surveys, due to its larger user base. The World Bank's MENA Data Lab set up the "World Bank MENA Pulse", through which internet surveys are administered. For example, ads are shown to specific sub-populations on Facebook, stratified by gender, age and geographic location to obtain a broad sample of Facebook users. This was used [to track the welfare implications of inflation](#) and to obtain information on racism and discrimination in MENA region. The food security survey found that in Morocco, cereals, fuel, transport, and fresh vegetables were most affected by rising prices, that price increases affected households in the middle part of the asset index distribution from the collected sample most, and that urban populations perceived the price increases to be stronger than rural respondents (except for cereals).

These types of survey are useful in response to emergencies, when there is a strong need for rapid data collection, and to obtain quick information on sentiments and current developments. Hoy et al. (2023) use data from 37,000 respondents in 12 middle-income countries to identify factors that drive or inhibit public support for energy subsidy reform. Respondents were drawn from a pre-registered pool, which allows for a fast collection of data and for fewer non-responses as respondents are familiar with the type of surveys, and are broadly representative of internet using population in each country. Another example is [Mobile Engage](#), a

mass text messaging system that was set up in Tajikistan and Uzbekistan for mass public outreach, but can also be used to collect information from respondents. As such, online and messaging-based surveys are a useful tool to reach large samples in a very swift manner and collect welfare proxy indicators. The downside, which limits its use as inputs to welfare estimation methods, is that samples will not be representative for the whole population. Moreover, not having contact with an enumerator can reduce the trust in a survey and removes the possibility to ask clarification questions.

Further Resources

Resources on data collection and welfare estimation using HFPS

- Several guidelines, best practices, and lessons learned have been published, which facilitate phone survey data collection, enhance, and sustain data quality, and reduce sampling biases (see for example, Himelein et al. 2020, Ambel et al. 2021, Brubaker et al. 2021, Gourlay et al. 2021). The [Poverty and Equity Global Practice \(GP\)](#) and the [Development Data Group \(DECDG\)](#), together with many other World Bank Global Practices and National Statistical Offices (NSOs), partnered to develop survey tools and technical guidelines to design and implement phone survey systems:
 - See [Brunckhorst et al. \(2023\)](#) for lessons-learned on HFPS during the COVID-19 pandemic.
 - See Himelein et al. (2020) for [guidelines on sampling for high-frequency phone surveys](#).
 - A World Bank [questionnaire template](#) with core and optional modules for HFPS, accompanied by an [overview](#) and a [manual](#), and [Guidelines on the implementation of Computer-assisted telephone interviewing \(CATI\) -based data collection](#).
 - The [World Bank Fragility, Conflict and Violence website](#) contains resources and links to datasets and collection guides.
- Literature on the estimation of poverty rates
 - [Monitoring Social and Economic Impacts of COVID-19 on Refugees in Uganda](#) using the SWIFT approach.
 - [Mahler et al. \(2022b\)](#) use information from HFPS on income gains or losses, combined with national accounts data, to estimate distributional impacts of the pandemic.

- Literature on other proxy welfare indicators:
 - [Abay et al. \(2023\)](#) estimate food security in Ethiopia.
 - [Brunckhorst et al. \(2023b\)](#) estimate the effects of COVID-19 on employment, self-reported welfare, food security, and asset depletion using HFPS.
 - [Kugler et al. \(2022\)](#) investigate changes in employment during the pandemic.
 - [Tabakis et al. \(2022\)](#) look at changes in self-reported welfare in countries affected by fragility, conflict, and violence during the pandemic.
 - See Table A6 in [Brunckhorst et al. \(2023a\)](#) for a summary on papers using HFPS.

Comparing survey costs

While the associated costs to implement a survey will vary by setting, some efforts have been made to compare the cost-effectiveness of different approaches. For example, Amaral et al. (2022) compare the cost-effectiveness of phone and messaging-based surveys in El Salvador, finding that survey-completion is 42 percentage points larger in phone surveys. In that case, higher implementation costs were more than outweighed, making phone surveys more cost-effective. Table 1 presents some estimated ranges for the cost per interview of implementing different types of surveys.

TABLE 1 Review of interview costs along different survey methods

| Survey type | Cost per interview | Country | Source |
|------------------------------|---------------------------------------|--------------|---|
| Traditional household survey | \$50-\$400, depending on the country | 18 countries | Kilic et al. (2017) |
| Phone surveys | \$2 to \$31, depending on the country | 38 countries | Brunckhorst et al. (2023b) |
| Community-based surveys | \$1.50 | Malawi | Decision Note for RFMS Malawi |
| Online surveys | \$2-\$3 through Facebook ads | MENA | The MENA Data Lab Knowledge Note Series |

| Survey type | Cost per interview | Country | Source |
|-----------------------|--|-------------|--|
| Message-based surveys | \$6.20 (\$21.20 adjusted by probability of completion) | El Salvador | Amaral et al. (2022) |
| SWIFT | \$15-\$25, if collected by phone, negligible cost if added to existing survey | | Yoshida & Aron (<i>forthcoming</i>) |
| SWIFT 2.0 | Can save between 5-25 percent of traditional survey, depending on how much cheaper it is to collect only correlates. | | Fuji and van der Weide (2020), Yoshida et al. (2022) |

Note: Costs can vary substantially depending on the country.

2.2 Geospatial Data

Main Characteristics and Examples

There is a wide range of geospatial data, with an even wider range of applications. Geospatial data refers to any data that has a spatial or geographic component—ranging from satellite images to household survey data that includes geographic coordinates. Geospatial data has long been in the toolkit of micro-economists, allowing them to analyze how welfare differs across a territory—for example, through poverty mapping. Advances in the quality of satellite imagery, as well as in the spatial and temporal resolution of remotely-sensed data,³⁷ have expanded the use cases for welfare mapping and monitoring.³⁸ Moreover, with cloud computing systems and machine learning algorithms, it is becoming easier to access and use vast amounts of geospatial data. Many remote sensing products are accessible in a pre-processed format, in which usable features have been extracted from raw imagery. While this section primarily discusses remotely-sensed geospatial data products, other types of geospatial data such as geo-tagged locations of households, infrastructure, or public services are also of great importance for welfare analysis.

³⁷ Remote sensing data are products that are collected from earth observation satellites or airplanes.

³⁸ In 2023, there are over 1000 earth observation satellites in orbit ([Union of Concerned Scientists Satellite Database \(2023\)](#)).

A large advantage of geospatial data products, such as nighttime lights data, is that poverty and welfare estimates can be produced at subnational levels and more frequent intervals. Therefore, geospatial data can be leveraged for all three purposes outlined in Figure 9. Relatedly, a new World Bank project is setting out to develop a method to predict welfare distributions for all country-year pairs for which no survey data is available based on PIP, WDI, and remote-sensing data. Over the past decade, the Poverty and Equity GP has increasingly tapped the potential of geospatial and remote-sensing data to fill gaps in other data sources. The main types of remote-sensing data used by the Poverty and Equity GP are presented below. Readers should keep in mind that there can be strong caveats using geospatial data, which are presented in the next sub-section.

Spatial data with pre-processed features

Pre-processed spatial data is accessible and easier to use than raw satellite imagery. Examples include vegetation indices, population density, nighttime lights, or proximity to relevant infrastructure, or social services. Instead of extracting information from raw imagery, these types of data allow for faster use, as extracted features are readily available. Newhouse (2023) finds that in the existing literature, models using interpretable features perform at least as well as those relying on raw imagery using convolutional neural networks. Several World Bank analyses use these types of data. As an example, the [Central African Republic Poverty Assessment 2023](#) uses geospatial variables such as indicators for urbanization or vegetation health to produce poverty maps.³⁹

Nighttime lights data is frequently used in economic studies as it offers a comparatively easy way to identify human activity. Nighttime lights data have been widely leveraged for a variety of purposes, such as measuring economic growth, urbanization, or population density (Zhang and Seto 2011, Henderson et al. 2012). Nighttime lights were also among the first applications of satellite data for welfare measurement (Elvidge et al. 2009, Chen and Nordhaus 2011). As nighttime lights data is available in close to real-time, it can be useful both as a proxy for welfare and as an input into imputation models as auxiliary data (Michalopoulos and Pappaianou 2018). There are several applications of proxying GDP growth (Beyer et al. 2022; Hu and Yao 2022; Martinez 2022) and measuring the impacts of shocks (Kocornik-Mina et al. 2020, Beyer et al. 2021) using nighttime lights. The

³⁹ In the absence of census data, geospatial data can complement other data to produce poverty maps, which is a useful tool to employ when subnational poverty data is of interest. (Lee and Braithwaite 2022; Corral et al. 2022; Newhouse 2023).

[Mozambique Poverty Assessment 2023](#) studies the effect of weather shocks on economic activity in urban areas, using nighttime lights data. However, several researchers have noted caveats for predicting poverty rates from nighttime lights data directly, which are discussed in more detail below.

For populations heavily dependent on agriculture, vegetation indices have been used as predictors for poverty.⁴⁰ The normalized difference vegetation index (NDVI) can accurately predict poverty and consumption rates in real time for rural and poor communities that rely heavily on agriculture (Tang et al. 2022). In the [Lake Chad Regional Economic Memorandum](#), NDVI was utilized to identify droughts and income shocks for agricultural societies, showing that anomalies in the NDVI index are associated with conflict onset. Vegetation yields can also complement other data for growth accounting in countries where data quality is low (Angrist et al. 2022).

Remote-sensing products are also the basis for early warning systems and can be used as leading indicators, for instance to identify potential hurricanes, droughts or flooding events. The World Bank is improving statistical models that can predict the future emergence of food crises, which may help to increase lead time for action (Andree et al. 2020). Gascoigne et al. (*forthcoming*) estimate high and significant welfare costs of droughts in Sub-Saharan Africa, which could be curtailed with early warning systems and preventative mechanisms. Cooper et al. (2019) use data on environmental factors, governance, and political instability alongside survey data to map locations with the largest vulnerability of child stunting due to drought. A promising new source of data is solar-induced chlorophyll fluorescence (SIF), which newer satellites can detect (Mohammed et al. 2019; McBride et al. 2022). SIF can detect vegetation stress at much earlier stages than the commonly used NDVI. Such variables can serve as leading indicators for upcoming crises. Remote sensing has also been used for early warning systems in other contexts, such as damage from earthquakes, natural catastrophes (see an example from [Central America](#)), or [global food security](#).

Other research shows the potential of this kind of early warning information to inform rapid policy response to protect household welfare. For example, Pople et al. (2021) analyze the use of data-driven forecasts of floods to trigger the release of anticipatory cash transfers in Bangladesh. They found that the anticipatory

⁴⁰ Burke & Lobell (2017) show that agricultural productivity can be measured accurately with satellite imagery.

transfers increased welfare by enabling households to take different decisions that altered the flood impacts at a critical juncture in time.

Daytime satellite imagery

Daytime satellite imagery reveals other types of information that may be relevant for estimating welfare, such as road quality, building size, roofing material, or agricultural land. Jean et al. (2016) combine nighttime lights with daytime satellite imagery and household surveys, which improves the accuracy of welfare predictions for areas without survey coverage. The authors argue that linear regressions of welfare on nighttime lights does not capture variations for poorer households, and therefore employ a transfer learning approach (see also the section on covariate-based imputation and annex 2 on machine learning methods), by using convolutional neural networks to associate daytime satellite images, such as roads, buildings or roof material, with nighttime lights data. A model is then trained on cluster-level welfare indicators from survey data, to make out-of-sample predictions in five countries in Sub-Saharan Africa. Watmough et al. (2019) predict household level welfare based on satellite imagery (for instance, of building and agricultural land size) with an accuracy of 45 percent for the whole sample, and of 62 percent for the poorest households in the sample.

Satellite imagery could also be used to infer the effectiveness of anti-poverty programs. Huang et al. (2021) measure changes in the quality of housing using a deep learning model. Comparing beneficiaries and control households of the randomized GiveDirectly program, the authors find that satellite imagery delivers results consistent with estimates from field surveys. However, the success of the evaluation hinges on the set up of the program: in this particular case households were targeted based on poor roofing quality, and were to upgrade their roofs when receiving cash transfers. Therefore, in cases when program targeting or evaluation is not based on easy-to-observe features, satellite imagery is less likely to be helpful in measuring impacts and program effectiveness.

Other spatial data reflecting economic activity

In many small island states, data can be very scarce. In the Pacific Islands, the Pacific Observatory initiative of the World Bank is experimenting with different sources to complement limited official statistics. For instance, a method developed by the IMF uses Automatic Identification System (AIS) data of merchandise

trade ships to estimate trade flows (Arslanalp et al. 2021).⁴¹ [A partnership between IMF and World Bank](#) uses this method to [nowcast economic activity](#). Verschuur et al. (2021) use global ship tracking data to estimate reductions in trade flows during the COVID-19 pandemic, distinguishing between ports and sectors and revealing that small island nations were hit hardest.

There are also other examples of how specific types of spatial data can proxy economic activity. For instance, gas flaring, which is correlated with oil production, can be used to measure economic activity in oil-producing countries such as Yemen (Debbich 2019) or of terrorist organizations (Do et al. 2018).⁴² Researchers have also tracked the number of cars in parking lots of retail stores to measure the impact of the COVID-19 pandemic or of natural disasters (Minetto et al. 2019). Such applications are, however, rare for low-income countries.

These examples showcase innovative ways of leveraging existing data for specific use cases. However, in most cases, the advantage is to monitor short-term changes in economic activity, rather than changes in individual or household welfare. Applications can be useful when the aim is to observe whether specific types of shocks, such as inflation or tariff changes, have an effect on consumption. Rarely can they be used to measure impacts on poverty or welfare for broader populations.

Sensors, drones, and other collection devices can also be deployed to generate high-frequency data. An example of such work by the Poverty and Equity GP is a study of the effects of air pollution on welfare (Baquie et al. 2023). The study collected real-time air pollution data from pollution monitors in Tbilisi, Georgia. Results show that air pollution leads to respiratory and mental health issues, and that poorer and less-educated households are more exposed to higher levels of pollution and employ fewer adaptive and protection measures.

Caveats for Using Geospatial Data

The following addresses some caveats that are particularly relevant for geospatial data applications in the context of welfare monitoring and provides links to further resources.

⁴¹ See also this application of [vessel tracking data to nowcast monthly trade of the Solomon Islands](#).

⁴² See also the World Bank's [Global Gas Flaring Tracker](#).

While satellite imagery is readily available globally and frequently, coverage can still be low in certain areas due to technical issues. For instance, in Tonga, only four days between November 2021 to February 2022 had clear satellite imagery because of clouds and low satellite revisit rates. Active sensors (such as Synthetic Aperture Radar, or SAR) can penetrate through clouds to produce day and night-time imagery, but only a small number of satellites are equipped with such sensors. Another option to deal with cloud coverage is to merge multiple images within a certain timeframe to create a cloud-free mosaic. However, when rapid data is needed, for instance, of a flooding during a monsoon season, cloud coverage may inhibit data availability. The usefulness of nighttime lights data in measuring the negative effects of natural hazards has been contested, even when methods to decrease volatility and noise have been applied, which is partly driven by the issue of cloud cover (Skoufias et al. 2021).

As nighttime lights are a good proxy for a broad range of variables, it can make distinguishing poverty and welfare from other factors difficult, and it does not predict time-series well (Chen and Nordhaus 2019; Asher et al. 2021). Furthermore, by itself, the accuracy of nighttime lights data to predict poverty rates is low in some regions (Engstrom et al. 2022). Nevertheless, Beyer et al. (2018) show that nightlights can be informative in measuring the impacts of specific shocks (such as natural disasters, conflict, and demonetization). That nighttime lights do not reflect welfare well in some regions can be driven by the agricultural sector, which is not well captured by nighttime lights (Beyer et al. 2018). Similar concerns have been raised for rural areas, which emit lower levels of nighttime lights (Chen and Nordhaus 2019; Gibson et al. 2020). Keola et al. (2015) complement nighttime lights data with land cover data to identify areas characterized by agricultural activity. Most studies using nighttime lights differ from those using other satellite imagery in that the former typically directly estimate elasticities in a linear framework, rather than predicting wealth levels based on survey data. Jean et al. (2016) and Yeh et al. (2020) point out that nighttime lights data alone does not distinguish well between poverty levels among lower income households.

An important disadvantage of daytime satellite imagery is that inputs reflect stock rather than flow indicators, making it difficult to track changes in welfare over time. Deep learning models based on satellite imagery and survey-based wealth estimations have been found to predict 50 percent of the variation of district-level changes in wealth over time (Yeh et al. 2020). Satellite imagery alone may therefore not be apt at capturing changes in poverty rates after economic shocks and for real-time monitoring.

A nascent literature is investigating biases and measurement errors in applications using remotely-sensed or other types of big data, and machine learning algorithms with limited ground-truth data. Comparing regression analysis based on ground-truth data and remotely-sensed data, Proctor et al. (2023) find that true parameters are often not within estimated 95 percent confidence intervals. Biases arise from classical measurement error and structured measurement error when remotely-sensed data is processed using machine learning methods. Multiple imputation can improve accuracy of predicted parameters and confidence intervals. Huber and Mayoral (2024) document similar non-classical measurement error when using nightlights as either the dependent or independent variable. Aiken, Rolf, and Blumenstock (2023) outline similar errors in applications of constructing satellite-based poverty maps. For instance, in many countries, machine learning models identify wealthy areas as those that are urban, and poor areas as those that are rural, hence limiting the ability to identify poverty within urban and rural areas. Recalibration methods can somewhat reduce biases. Rolf (2023) provides an overview of model evaluation for geospatial machine learning.

There is always a challenge in merging and aggregating geospatial and survey data. To be able to directly extract geospatial variables to household survey data, the latter needs to include geocoordinates. This is not always the case, especially as accessing household locations can violate privacy rights. Using different levels of aggregation, or aggregating geospatial indices to larger levels so that they are economically meaningful can diminish its usefulness and lead to biases (this is called the modifiable areal unit problem (Fotheringham and Wong (1991))). Averaging vegetation indices at state levels may conflate useful local information in household locations. Moreover, as Tang et al. (2022) show, NDVI values can be predictive for welfare in rural areas, but less so in urban regions. Similarly, calculating population densities by dividing population numbers by the total area of larger administrative levels can cause mismeasurement if populations are densely populated in only a few areas within a state. However, Hall et al. (2023) do not find a statistically significant relationship between the spatial resolution of satellite imagery and model performance.

Many geospatial data products are the result of predictions and modelling. Gridded population data, for instance, is developed by modeling the distribution of population with other characteristics such as built-up areas (GHSPop) or land cover, nighttime lights, infrastructure, and other spatial inputs (WorldPop). Such modeled inputs can increase measurement error and uncertainty in predictions.

Lessons Learnt and Resources

Geospatial data has opened many opportunities for welfare monitoring, but, with current models, remotely-sensed data is best used as a complement to other data sources. The (near) universal availability of remotely-sensed data enables monitoring and measurement in regions where other data is sparse or non-existent. Burke et al. (2021) emphasize that geospatial data should be seen as enhancing rather than replacing survey data. Feeding geospatial data into imputation models combined with household survey data improves welfare estimates. Geospatial data can, for instance, be used to improve survey-to-survey imputation models (Dang et al. 2023). For the purposes of real-time monitoring, it needs to be kept in mind that in many applications remotely-sensed data works well in cross-sectional analysis, but less so in monitoring changes over time.

Remote-sensing products can also be leveraged for project and damage monitoring. For instance, satellite data can be used to detect and track deforestation rates, crop yields, water quality or road or electricity infrastructure construction (see [Remote sensing guide for practitioners](#)). The effectiveness of electrification projects has been monitored using nighttime lights data in [Senegal and Mali](#). A project in [Myanmar](#) will assess the damage of a cyclone, which struck mid-May 2023. A project in the [Ukraine](#) is using spatial pollution measurement to identify the environmental hazards from the conflict. In the [Democratic Republic of Congo](#), geospatial tools are used to measure the extent of deforestation.⁴³ When possible, using such data can significantly reduce costs of monitoring and evaluation.

Some resources of geospatial data include:

- Access the World Bank repository [Open Nighttime Light Tutorials](#) for information on processing, analyzing, and developing data products from nighttime lights.
- [This World Bank repository](#) contains the Light Every Night (LEN) dataset, which consists of the complete archive of all nighttime lights imagery captured each night over the last three decades.⁴⁴

⁴³ Changes in forest coverage have also been used to study the impact of the Rohingya displacement in Cox's Bazar on deforestation and in-migration (Dampa et al. 2022). When such correlations exist, changes in observable geospatial indicators can be leveraged to capture hard-to-measure socioeconomic developments.

⁴⁴ The underlying data is sourced from the NOAA/NCEI archive (DMSP-OLS with data from 1992-2017, and the VIIRS-DNB with data spanning 2012-2020). The World Bank worked in collaboration with NOAA/NCEI and the University of Michigan to publish the archive as an Analysis-Ready Dataset. The

- World Bank efforts are making [nighttime lights data more accessible](#) over longer periods of time and at higher resolution. Newer data from VIIRS (the Visible Infrared Imaging Radiometer Suite) has been found to be [more accurate](#) and fine grained than the widely used DMSP (Defense Meteorological Satellite Program) data (Gibson et al. 2021).⁴⁵
- The [Geospatial Technology and Information for Development](#) includes important takeaways and resources from the World Bank's work in this area.
- See the [Remote sensing guide for practitioners](#) guide for resources and example projects of using remote-sensing data.
- [The ITS Geolab](#) interface and repository enables access and overlay various spatial data.
- [The COVID-19 economic data toolkit](#) includes resources on nighttime lights data.
- Within the World Bank, the WBG-internal Geospatial Operations Support Team (GOST) provides remote-sensing support for select operations. The [Geospatial Platform](#) also includes datasets and other various resources.
- There is also a WB guide for practitioners on Unmanned Aerial Vehicle Operation regions (see this [practitioner's guide](#)).
- [Stewart et al. \(forthcoming\)](#) propose a standardized five-step workflow for geospatial analysis in situations of forced displacement.
- [DIME WIKI](#) on Remote sensing, including a guide to use Google Earth Engine.
- For remotely-sensed data sources, see, for instance, [Landsat](#) or [Modis](#).
- [Earth Genome](#) has developed a tool that uses AI to identify features on satellite imagery. Project proposals can be submitted through [Development Data Partnerships](#).

LEN archive, which now spans nearly 250 terabytes and is published under the World Bank's open data license, is available on the AWS open public dataset program.

⁴⁵ When using raw nighttime lights data, researchers should reduce background noise following procedures as laid out by Beyer et al. (2018) or Elvidge et al. (2017).

- [MOSAIKS](#) uses a machine learning approach to convert a very large set of satellite imagery into tabular data. This is used e.g. by [Marty & Duhaut \(2024\)](#). More information is on this [blog](#).

Example of literature using geospatial data for welfare estimation:

- Predicting welfare with spatial data with pre-processed features
 - [Van der Weide et al. \(2023\)](#) find that combining a range of geospatial variables with survey data can yield an accurate picture of the geography of poverty, but not of accurate estimates of poverty for specific regions.
 - [Merfeld and Newhouse \(2023\)](#) use machine learning methods to estimate poverty rates or asset indices in four countries with NDVI, population, temperature, distance to city, land cover, and other data.
 - Poverty Assessments, for example the [Central African Republic Poverty Assessment](#), use geospatial indicators, including urbanization or vegetation health, to produce poverty maps. The [Mozambique Poverty Assessment 2023](#) uses nighttime lights data to measure weather shocks in urban areas.
 - [Tang et al. \(2022\)](#) (employing convolutional neuro networks) and the [Lake Chad Regional Economic Memorandum \(using linear regression models\)](#) predict poverty rates and income shocks using vegetation indices. Estimating consumption rates based solely on vegetation indices is likely to perform well only in rural and agricultural areas.
 - [Andree et al. \(2020\)](#) predict food crises with vegetation indices using random forest models.
- Predicting welfare with raw satellite imagery
 - [Marty & Duhaut \(2024\)](#) predict asset-based welfare from DHS data using different raw and preprocessed geospatial data sources, including nighttime lights, daytime imagery and vegetation indices, as well as four different machine learning algorithms.
 - [Jean et al. \(2016\)](#) and [Yeh et al. \(2020\)](#) use deep learning methods on satellite imagery and household surveys to predict welfare in local areas in African countries.
 - [Lee and Braithwaite \(2022\)](#) employ an iterative process of predicting wealth classes using extreme gradient boosting based on labeled spatial features and using a CNN model with unlabeled satellite images.
 - [Watmough et al. \(2019\)](#) extract features from satellite imagery at very fine resolution, to identify land-use and predict household-level wealth.

- [Huang et al. \(2021\)](#) predict household-level wealth based on housing quality, derived from satellite imagery.
- Survey-to-survey imputation
 - [Dang et al. \(2023\)](#) improve the accuracy of survey-to-survey imputation using linear regression models by adding geospatial variables in the imputation process.

2.3 Digital Trace Data

Main Characteristics and Examples

Digital trace data encompasses information or data that is generated and collected from digital activities. For instance, mobile network operators have large databases on customers, which are termed Call Detail Record (CDR) data. Such data includes information on phone usage behavior, such as the number of calls made and received, call duration, call destination, as well as the location of mobile phone users, the type of phone used, and their payment history. CDR data can be leveraged to infer information about people such as their mobility patterns, consumption behavior, and social network connections. Similarly, social media platforms host vast amounts of data on user locations, frequency and type of usage, and social connections, that in some cases can be utilized for research purposes.

Phone and social media data, in combination with other data sources, has been used for welfare and poverty estimations, and there have been some successful applications. Most applications using CDR or social media data entail machine learning methods, as datasets are large, complex and noisy.⁴⁶ Supervised machine learning techniques, trained on survey and mobile phone usage data, including frequency, length and incoming and outgoing calls, have for instance performed well in predicting district-level poverty rates in Rwanda (Blumenstock, Cadamuro and On 2015). Similar results have been found for predicting poverty in Afghanistan (Blumenstock 2018), in Guatemala, though with more accuracy in urban than in rural areas (Hernandez et al. 2017) for predicting multidimensional poverty in Senegal in combination with environmental data (Pokhriyal and Jacques 2017) or for predicting food insecurity in a central African country (Decuyper et al. 2014). Steele et al.

⁴⁶ As CDR in most cases includes locational information, it also constitutes a type of geospatial data.

(2017) also show that combining CDR and other geospatial data, including nighttime lights and vegetation indices, can predict welfare well in Bangladesh. The authors show that CDR improves estimates especially in urban areas, which highlights the potential of complementing data sources that perform better in different settings.

To improve the targeting of humanitarian assistance in environments without survey data, researchers have used machine learning algorithms, trained on traditional survey data, to recognize patterns of poverty in mobile phone data, for instance in Togo (Aiken et al. 2022) or Afghanistan (Aiken et al. 2020). Not all studies have been successful in using CDR for poverty assessment or impact evaluation, however. Barriga-Cabanillas et al. (2022) conducted a study to evaluate the accuracy of a machine learning model fed on CDR data in estimating the effects of a cash-transfer program in Haiti, compared to results obtained from a regression discontinuity design with survey data. The machine learning model fails to identify beneficiaries and does not show statistically significant effects, whereas the regression discontinuity model yields large and positive impacts.

Digital trace data goes beyond CDR data to include a vast range of sources such as web search queries or social media activity. Researchers have combined connectivity data from Facebook with data from satellites, mobile phone networks, and topographic maps to construct a relative wealth index (RWI) for of all 135 low and middle-income countries at 2.4 km resolution (Chi et al. 2021). Their model is trained on DHS data from 56 countries using deep learning algorithms. The RWI has been used by Poverty and Equity GP staff for monitoring purposes or analyzing the distribution of poverty. The number of social media users can predict a wealth index from DHS data with an accuracy of about 60 percent in India and Philippines (Fatehkia et al. 2020).

Digital trace data can therefore provide signals for economic and social activity, before these are captured by surveys. Social media data can be used to infer welfare correlates such as unemployment trends (D'Amuri and Marcucci 2017) or subjective well-being (Voukelatou et al. 2021). Mobility patterns from CDR can provide important information on migration flows (Beine et al. 2019; Fraiberger et al. 2020; Blanchard et al. 2021; Olivieri et al. 2022;). Data from Facebook has been used to [track the mobility of refugees from the Ukraine or Syria](#), the outflow of migrants from Venezuela, and [population movements due to COVID-19](#) (Palotti et al. 2020).⁴⁷

⁴⁷ Chetty et al. (2022a, 2022b) make use of friendship connection data from Facebook to explore patterns of social networks and determinants of socioeconomic mobility.

Caveats for Using Digital Trace Data

The use of digital trace data is a new field and applications to welfare and project monitoring are still being tested. While there are some promising attempts to leverage big data for welfare monitoring, in many cases, validity and scalability need to be verified. The accuracy of identifying poor populations based on mobile phone data may be sufficient in situations where no other data is available, but likely not for routinely collected poverty statistics (Aiken et al. 2021).⁴⁸ Furthermore, phone data may not work well in detecting idiosyncratic shocks, or vulnerabilities, or changes in flow, rather than stock variables (Blumenstock 2018, Barriga-Cabanillas et al. 2022). More evidence is needed to shed light on whether changes in mobile phone usage reflect broader changes in welfare and income, both immediately after shocks and in the longer term. Model stability issues also hold for applications of CDR data (Lazer et al. 2015, Blumenstock 2018).

As discussed above for the case of high-frequency phone surveys, not all individuals possess or use mobile phones, leading to substantial coverage bias. While there are methods to readjust sampling weights, more methodological research is necessary on how such biases can be mitigated when using big data. The study by Barriga-Cabanillas et al. (2022) in Haiti demonstrates that using CDR data is not useful in all settings. When phone ownership is low, and when populations are relatively homogenous, the data may not suffice to detect significant differences between more or less poor households. These issues are even larger for data from social media platforms. Lastly, in the study by Steele et al. (2017), the authors compare the predictive power of CDR and other geospatial data on different measures of welfare. While they find that an asset index derived from DHS data can be predicted well, they find lesser correlations with other outcomes such as income or consumption. Therefore, as is the case for many geospatial data, stock variables appear to be easier to predict than flow variables.

Accessing any type of personal data must come with careful considerations on data privacy and ethics. Digital trace data includes sensitive information. When using such data—in concordance with data privacy regulations and agreements—researchers and practitioners must be cautious in its storage, sharing, and in the presentation of results. Furthermore, algorithms trained on a specific source of data may be exposed to certain biases that could result in discriminative predictions, linked to measurement errors or sampling biases. Such biases are difficult to uncover, warranting caution to users. Moreover, in most

⁴⁸ In their study in Togo, Aiken et al. (2021) report an accuracy of around 70 percent.

countries, there are several mobile phone operators. For a nationally representative sample, in most cases data from multiple operators will be necessary, which can be difficult to obtain.

Lessons Learnt and Resources

Digital trace data can be used to complement other data sources to allow for more fine-grained mapping of poverty or targeting of vulnerable populations. Using digital trace data for poverty and welfare monitoring is still a relatively new field, with many new approaches currently being tested. The main advantage of big data so far is to fill in data gaps when other sources are not available or not recent, such as for targeting the poor in emergency situations (Aiken et al. 2021), or as a complement to other data sources to improve estimates and granularity (Chi et al. 2021). Digital trace data opens avenues for new applications, such as using Google search data to detect the onsets of crises, an example being the COVID-19 pandemic (Ten et al. 2022).

By itself, it is likely not yet feasible to obtain accurate welfare and poverty estimates from digital trace data for a wider range of countries in different contexts. While there have been promising applications for leveraging CDR data and machine learning techniques to predict poverty rates, these procedures have not been successful in specific country contexts. Many methodological issues in using such data have yet to be addressed. While it can be very useful for particular use cases, such as project monitoring or tracking mobility patterns, it may not be feasible to estimate poverty and welfare globally.

Some available resources can be found here:

- [Connectivity Mapping](#): A guide to access data from advertisement platforms from social networks via API, and to use the data to map number of users using the particular platform, as well as user characteristics (such as demographics, education, and job experience).
- [Data in Action](#): A repository with materials to guide the process of designing and delivering data products in the context of international development.
- [Twitter for Economic Monitoring](#): A tool and training material to generate indicators from Twitter data with a focus on unemployment, public sentiment, and

misinformation during the COVID-19 pandemic, with practical examples from Brazil, Mexico, and Pakistan.⁴⁹

- [Handbook on the Use of Mobile Phone Data for Official Statistics](#): A guide on mobile phone data by the UN Global Working Group on Big Data for Official Statistics
- World Bank staff has developed a technical note on how to use human mobility data to inform urban and disaster risk-management engagements. The note offers guidance on how to use mobility data, for instance, from Facebook, Cuebiq, and Mapbox for a variety of purposes, including caveats and limitations (Jones et al. *forthcoming*).
- [rsocialwatcher](#) is a R package developed by Marty and Duhaut (2024), which facilitates querying Facebook Marketing API.

Example of literature using digital trace data for welfare estimation:

- Gradient Boosting or Random Forest for welfare prediction
 - [Blumenstock, Cadamuro, and On \(2015\)](#) and [Blumenstock \(2018\)](#) use gradient boosting on CDR data to predict welfare rates in Rwanda and Afghanistan.
 - [Aiken et al. \(2022\)](#) and [Aiken et al. \(2020\)](#) use the same approach to impute consumption and wealth of households in small areas in Togo and Afghanistan, with the goal of improving the targeting of humanitarian aid.
 - [Hernandez et al. \(2017\)](#) employ supervised machine learning techniques on CDR data from Guatemala to predict subnational poverty rates, with higher accuracy in urban than rural areas.
 - [Barriga-Cabanillas et al. \(2022\)](#) use different supervised machine learning models to impute food security from mobile phone data, but document poor performance in the predictions.
- Bayesian Models for welfare prediction
 - [Steele et al. \(2017\)](#) use Bayesian Geospatial Models on CDR and environmental data to predict non-monetary and monetary poverty in Bangladesh.

⁴⁹ World Bank staff has access to a continuous stream of 10 percent of all tweets.

- [Pokhriyal and Jacques \(2017\)](#) predict multidimensional poverty rates in Senegal using environmental and CDR data and Gaussian Process regressions (a Bayesian learning technique).

2.4 Administrative Data

Main Characteristics and Examples

Administrative data refers to datasets that are not collected by surveying individuals or households and do not have the primary goal of providing information for research. In many cases, administrative data falls under the broader category of big data. It encompasses a variety of official public or private sector records, such as financial transactions, tax statements, loan payments, birth records, records of applicants and/or participants in social or work programs, or school enrolment.

The advantages of administrative data are that it typically is very detailed and covers the whole universe of a particular group. Since the data is already “on record”, utilizing it incurs a relatively low cost and requires less time than collecting data. Administrative data can also help to overcome inherent issues with surveys, such as misreporting. Meyer and Mittag (2019) find that survey data in the U.S. understates incomes of poor households, distorts program targeting, and severely understates the effectiveness of anti-poverty programs. Administrative data, in this case of payments from welfare programs, including data on recipient addresses, date of payment and payment amounts, help to correct these numbers. In some countries, administrative data have been heavily used by researchers, such as in Denmark, Norway, Finland, Sweden, or the United Kingdom.

Administrative data is particularly useful when it can provide information on entities for which collecting data can be difficult, or for tracking subjects or items over long periods of time, which would be costly and difficult with survey data. For instance, it has made it possible to follow income dynamics of immigrants in Canada (Picot et al. 2019) or of technical and vocational training graduates in Saudi Arabia (Rivera et al. 2022). Tax record data has been leveraged to study productivity or employment outcomes in firms, for instance in Chile (Albagli et al. 2023) or South Africa (Pieterse et al. 2018). Beyer et al. (2021) use daily electricity consumption records from a government-owned utility enterprise in India to analyze the impact of COVID-19 over time. Cuesta and Chagalj (2019) find that including administrative macroeconomic data in microsimulations significantly improves poverty estimates in Nicaragua.

Specific types of administrative data have the potential to enhance welfare and poverty measurement. Barcode scanner data could, for instance, be used for real-time inflation measurement (Dubois et al. 2022).⁵⁰ Beck and Jaravel (2020) introduce a harmonized barcode-level dataset on expenditure and prices from 34 countries, including some middle-income countries. The authors conclude that standard price indices are biased downward for middle-income and more populated countries, which entails international inequality being underestimated and GDP per capita measures being over-estimated.⁵¹ For many countries, however, especially those that are already data-deprived, such data does not exist.

Tax records encompass direct information about labor and capital income, which can be utilized to measure poverty and income distributions directly. Depending on when such data is published or made available by government entities, it can fill gaps between survey waves. Larrimore et al. (2022) utilize income tax and unemployment benefits data to estimate changes in earnings during the COVID-19 pandemic in the U.S. Morgandi et al. (2022) combine several public sector administrative data sources to identify and study the working poor in Brazil. Also in Brazil, Alvarez et al. (2018) match employer and employee records to analyze earning inequality. Brazil is, however, one of few countries that make public administrative records available frequently. Another source of data to analyze immediate effects of shocks are job portals. Marinescu et al. (2020) and Forsythe et al. (2020) study how the number of job postings and applications changed after the start of the pandemic. Tax income data may furthermore advance measurement of top incomes: tax and survey data display similar income metrics along the income distribution, with the exception of the top 1 percent, for which survey data is not accurate, mainly due to non-labor income (Yonzan et al. 2021).

Other types of administrative data could also prove useful for welfare monitoring. For example, social registries (particularly dynamic social registries that allow for open and continuous registration) often contain detailed information on poor and low-income households. If these databases are of sufficient quality, have sufficient coverage, and are updated sufficiently regularly, they could provide valuable insights on the evolution of household welfare. Similarly, other types of administrative data linked to social programs (such as the ongoing collection of data on children’s school attendance in the context of monitoring

⁵⁰ Jaravel and O’Connell (2020) show inflation spikes during Covid-19 lockdowns using product-level scanner data from the United Kingdom.

⁵¹ Van der Weide et al. (2018) use house-price data to demonstrate that inequality is underestimated in Egypt.

conditions for cash-transfer programs) could provide timely information on other dimensions of household welfare. There are many types of administrative data whose value remains largely untapped for the purposes of high-frequency welfare monitoring.

Caveats for Administrative Data for Real-time Welfare Monitoring

The main, and obvious, reason why there are few use cases of welfare monitoring using administrative data is that it can be hard to come by, especially in lower-income countries. Collection and storage of vast amounts of data requires advanced infrastructure and capacity, which in many countries remains a challenge.⁵² Moreover, in many countries, high levels of informality negate the usefulness of administrative tax data. Some low-income countries, including Rwanda and Uganda, as well as some Pacific Islands countries make use of VAT data to compile GDP, monthly production indices, or business registries (Rivas and Crowley 2018).

Furthermore, it often includes sensitive data, such as personal details and information on addresses, and is therefore not accessible in many cases without strict confidentiality agreements and protection protocols. While administrative data covers the whole universe of a particular group, it can miss important populations, which are of particular interest for poverty estimates. For instance, income tax records can miss those not in the labor force. Certainly, being present or missing in an administrative dataset is not random, and correcting for such biases requires information on the whole population. Furthermore, there are concerns for accuracy of administrative data in some cases, which is a key reason why surveys are conducted in the first place (World Bank 2021b). Administrative data from private sources are a particularly selected sample.

Depending on the type of data, administrative records are only published with a time lag or refer to a past period, such as tax statements. Bachas et al. (2020) use tax-record data from firms of 10 low and middle-income countries to simulate effects of lockdown measures on sales and exit rates, but cannot estimate these rates directly as current data was not yet available. The direct use for real-time monitoring can therefore be limited.

⁵² As an example, complete civil and vital statistics systems exist in no low-income country, in 22% of lower-middle income, 51% of upper-middle income and 95% of high-income countries (World Bank 2021b).

Lessons Learnt and Resources

Administrative data can be a “gold mine” for specific use cases. Compared to other data sources, it is however much more difficult and less straightforward to find and gain access to suitable datasets. For specific use cases, it may be worthwhile for researchers and practitioners to assess the possibility and potential of using administrative sources. With statistical offices and data infrastructure in lower-income countries improving, the availability and accessibility of such datasets is likely going to increase in the future. Partnerships with the private sector could create new opportunities of using administrative data for welfare monitoring.

Resources:

- This [Handbook on Using Administrative Data for Research and Evidence-based Policy](#) by Shawn Cole, Iqbal Dhaliwal, Anja Sautmann, and Lars Vilhuber introduces readers to using administrative data and contains several case studies.
- See [this J-PAL guide](#) on using administrative data.
- [Vavra \(2021\)](#) provides an overview of efforts using administrative data to track real-time effects of the COVID-19 pandemic, though it is confined to the U.S.

3.

Moving Forward: *Identifying Areas for Advancement*

This typology offers a roadmap to some of the key approaches that can be leveraged to monitor welfare with greater frequency. It seeks to bring this work together in one place, compile a core set of resources, and reflect on the lessons learned from the Poverty and Equity GP's efforts in this area so far. While we have come a long way in advancing our knowledge on approaches for real-time monitoring, there are still many areas where questions remain and further work is needed. The analysis in this typology points toward the following specific areas as key areas for further exploration:

Area 1: Comparative Analysis of Nowcasting/Imputation Methods for Poverty Updates: Examining and contrasting various nowcasting and imputation methods for updating poverty is essential. Further exploration involves reviewing evidence and implementing new analysis to compare their performance within the same context and data ecosystem. For instance, in the LAC region, countries like Paraguay could provide insights through simultaneous implementation of nowcasting based on GDP, Adept, and survey-to-survey imputation. Similarly, in West Africa, analysis could expand by building upon panel surveys and existing nowcasting initiatives. Other potential cases could include Nepal and Bangladesh, where the data ecosystem may allow comparing various options.

Area 2: Leveraging Geospatial and other Big Data Sources: Access to these data sources is becoming cheaper and easier, and methods of analyzing vast amounts of data are improving. There have been promising examples of using geospatial and CDR data for poverty prediction. However, predictions that capture changes over time are still scarce, and oftentimes do not have high predictive power. This is an area where future research is much needed. Advancing on real-time monitoring using non-survey data will be key to nowcast poverty rates in data-scarce environments.

Area 3: Leading Indicators for Early Detection: Identifying and monitoring leading indicators capable of signaling potential changes in welfare is another area where more work is needed. A pilot initiative can begin by reviewing evidence on leading indicators useful for tracking and identifying areas or populations at risk of falling into poverty, followed by exploring efficient ways of collecting these indicators. Examples include inflation forecasts and climate indicators, such as droughts and flood forecasts.

Area 4: Clarifying the Role of Phone Surveys in Poverty Monitoring: Providing clarity and guidance on the conditions where phone surveys can be effective in monitoring welfare is necessary. This work involves a comprehensive review of evidence, identification of minimum conditions, synthesis and testing of survey reweighting methods, and identification of key variables essential for meaningful welfare monitoring.

Area 5: Enhancing Usability of Baseline/Auxiliary Household Surveys: Reviewing and summarizing recommendations to modernize questionnaires, ensuring comparability of core questions, incorporating geo-referencing, and providing re-contact information are crucial for enhancing the usability of baseline and auxiliary household surveys for monitoring purposes. Additionally, further work may include maximizing the utility of administrative data integration for welfare monitoring and devising strategies for monitoring welfare conditions in challenging settings, such as small island states and regions facing fragility or active conflicts. Despite not being the optimal tool for measuring poverty, labor force surveys can serve as an important source of auxiliary data due to their more frequent collection compared to budget surveys. Their representativeness and inclusion of income variables make them appealing instruments. Analyzing the value of using these surveys for poverty monitoring could be implemented in East Africa and MENA, where interest has been expressed.

References

- Abate, Gashaw T., Alan De Brauw, Kalle Hirvonen, and Abdulazize Wolle. “Measuring consumption over the phone: Evidence from a survey experiment in urban Ethiopia.” *Journal of Development Economics* 161 (2023): 103026.
- Abay, Kibrom A., Guush Berhane, John F. Hoddinott, and Kibrom Tafere. “Assessing response fatigue in phone surveys: Experimental evidence on dietary diversity in Ethiopia.” Vol. 2017. Intl Food Policy Res Inst, 2021.
- Abdul Rahman, Mariah, Nor Samsiah Sani, Rusnita Hamdan, Zulaiha Ali Othman, and Azuraliza Abu Bakar. “A clustering approach to identify multidimensional poverty indicators for the bottom 40 percent group.” *PloS one* 16, no. 8 (2021): e0255312.
- Albagli, Elías, Alejandra Chovar, Emiliano Luttini, Carlos Madeira, Alberto Naudon, and Matias Tapia. “Labor market flows: Evidence for Chile using microdata from administrative tax records.” *Latin American Journal of Central Banking* 4, no. 4 (2023): 100102.
- Alix-Garcia, Jennifer, and Daniel Millimet. “Remotely Incorrect? Accounting for Nonclassical Measurement Error in Satellite Data on Deforestation.” (2023).
- Aiken, Emily L., Guadalupe Bedoya, Aidan Coville, and Joshua E. Blumenstock. “Targeting development aid with machine learning and mobile phone data: Evidence from an anti-poverty intervention in Afghanistan.” In *Proceedings of the 3rd ACM SIGCAS Conference on Computing and Sustainable Societies*, pp. 310-311. 2020.

- Aiken, E., Bellue, S., Karlan, D., Udry, C. and Blumenstock, J.E., 2022. “Machine learning and phone data can improve targeting of humanitarian aid.” *Nature*, 603(7903), pp.864-870.
- Aiken, Emily, Esther Rolf, and Joshua Blumenstock. “Fairness and representation in satellite-based poverty maps: Evidence of urban-rural disparities and their impacts on downstream policy.” arXiv preprint arXiv:2305.01783 (2023).
- Alvaredo, Facundo, and Leonardo Gasparini. “Recent trends in inequality and poverty in developing countries.” *Handbook of income distribution 2* (2015): 697-805.
- Alvarez, Jorge, Felipe Benguria, Niklas Engbom, and Christian Moser. “Firms and the decline in earnings inequality in Brazil.” *American Economic Journal: Macroeconomics* 10, no. 1 (2018): 149-189.
- Amaral, Sofia, Lelys Dinarte, Patricio Dominguez-Rivera, Santiago M. Perez-Vincent, and Steffanny Romero. “Talk or text? Evaluating response rates by remote survey method during COVID-19.” (2022).
- Ambel, Alemayehu, Kevin McGee, and Asmelash Tsegay. “Reducing Bias in Phone Survey Samples: Effectiveness of Reweighting Techniques Using Face-to-Face Surveys as Frames.” *Policy Research Working Papers* (2021).
- Angrist, Noam, Pinelopi Koujianou Goldberg, and Dean Jolliffe. “Why is growth in developing countries so hard to measure?.” *Journal of Economic Perspectives* 35, no. 3 (2021): 215-242.
- Andree, Bo Pieter Johannes; Chamorro, Andres; Kraay, Aart; Spencer, Phoebe; Wang, Dieter. “Predicting Food Crises”. *Policy Research Working Paper*; No. 9412. © World Bank, Washington, DC (2020).
- Andree, Bo Pieter Johannes. “Machine Learning Guided Outlook of Global Food Insecurity Consistent with Macroeconomic Forecasts.” No. 10202. The World Bank, 2022.
- Arslanalp, Mr Serkan, Mr Robin Koepke, and Jasper Verschuur. “Tracking Trade from Space: An Application to Pacific Island Countries.” *International Monetary Fund*, 2021.

- Asher, Sam, Tobias Lunt, Ryu Matsuura, and Paul Novosad. "Development research at high geographic resolution: an analysis of night-lights, firms, and poverty in India using the shrug open data platform." *The World Bank Economic Review* 35, no. 4 (2021): 845-871.
- Auffhammer, Maximilian. "Quantifying economic damages from climate change." *Journal of Economic Perspectives* 32, no. 4 (2018): 33-52.
- Balashankar, Ananth, Lakshminarayanan Subramanian, and Samuel P. Fraiberger. "Predicting food crises using news streams." *Science Advances* 9, no. 9 (2023): eabm3449.
- Bañbura, Marta, Domenico Giannone, Michele Modugno, and Lucrezia Reichlin. "Now-casting and the real-time data flow." In *Handbook of economic forecasting*, vol. 2, pp. 195-237. Elsevier, 2013.
- Baquie, Sandra, Patrick A Behrer, Xinming Du, Alan Fuchs, and Natsuko K Nozaki. "Poverty and Distributional Consequences of Air Pollution in Tbilisi." Washington, D.C.: World Bank Group (2023).
- Barriga-Cabanillas, Oscar, Joshua E. Blumenstock, Travis J. Lybbert, and Daniel Putman. "Digital Breadcrumbs and Dietary Diversity: Testing the Limits of Cell Phone Metadata in Poverty and Impact Assessment." World Bank (2022).
- Beck, Günter W., and Xavier Jaravel. "Prices and global inequality: new evidence from worldwide scanner data." Available at SSRN 3671980 (2020).
- Bellon, Matthieu, Era Dabla-Norris, Salma Khalid, and Frederico Lima. "Digitalization to improve tax compliance: Evidence from VAT e-Invoicing in Peru." *Journal of Public Economics* 210 (2022): 104661.
- Beltramo, Theresa, Hai-Anh Dang, Ibrahima Sarr, and Paolo Verme. "Estimating poverty among refugee populations: A cross-survey imputation exercise for Chad." World Bank Policy Research Working Paper 9222 (2020).
- Bethlehem, Jelke. "Selection bias in web surveys." *International statistical review* 78, no. 2 (2010): 161-188.
- Bergstrom, Katy. "The role of inequality for poverty reduction." World Bank Policy Research Working Paper 7969 (2020).

- Beyer, Robert, Esha Chhabra, Virgilio Galdo, and Martin Rama. “Measuring districts’ monthly economic activity from outer space.” World Bank Policy Research Working Paper 8523 (2018).
- Beyer, Robert CM, Sebastian Franco-Bedoya, and Virgilio Galdo. “Examining the economic impact of COVID-19 in India through daily electricity consumption and nighttime light intensity.” *World Development* 140 (2021): 105287.
- Beyer, Robert CM, Yingyao Hu, and Jiaxiong Yao. Measuring quarterly economic growth from outer space. No. 9893. World Bank Policy Research Working (2022).
- Blanchard, P., Gollin, D., & Kirchberger, M. Perpetual Motion: High-Frequency Human Mobility in Three African Countries (No. tep0823). Trinity College Dublin, Department of Economics (2023).
- Blumenstock, Joshua, Gabriel Cadamuro, and Robert On. “Predicting poverty and wealth from mobile phone metadata.” *Science* 350, no. 6264 (2015): 1073-1076.
- Blumenstock, Joshua E. “Estimating economic characteristics with phone data.” In *AEA papers and proceedings*, vol. 108, pp. 72-76. 2014 Broadway, Suite 305, Nashville, TN 37203: American Economic Association, 2018.
- Bourguignon, François. “The growth elasticity of poverty reduction: explaining heterogeneity across countries and time periods.” *Inequality and growth: Theory and policy implications* 1, no. 1 (2003).
- Bourguignon, François. The poverty-growth-inequality triangle. No. 125. Working paper, 2004.
- Bourguignon, François, and Amedeo Spadaro. “Microsimulation as a tool for evaluating redistribution policies.” *The Journal of Economic Inequality* 4 (2006): 77-106.
- Bourguignon, François, and Maurizio Bussolo. “Income distribution in computable general equilibrium modeling.” *Handbook of computable general equilibrium modeling* 1 (2013): 1383-1437.
- Braley, Alia, Samuel P. Fraiberger, and Emcet O. Taş. “Using Twitter to Evaluate the Perception of Service Delivery in Data-Poor Environments.” Available at SSRN 4029916 (2021).

- Browne, Chris, David S. Matteson, Linden McBride, Leiqiu Hu, Yanyan Liu, Ying Sun, Jiaming Wen, and Christopher B. Barrett. "Multivariate random forest prediction of poverty and malnutrition prevalence." *PloS one* 16, no. 9 (2021): e0255519.
- Brubaker, Joshua, Talip Kilic, and Philip Wollburg. "Representativeness of individual-level data in COVID-19 phone surveys: Findings from Sub-Saharan Africa." *PloS One* 16, no. 11 (2021): e0258877.
- Bruederle, Anna, and Roland Hodler. "Nighttime lights as a proxy for human development at the local level." *PloS one* 13, no. 9 (2018): e0202231.
- Brunckhorst, Ben, Alexandru Cojocaru, Ruth Hill, Yeon Soo Kim, and Maurice Kugler. "Long COVID. The Evolution of Household Welfare in Developing Countries during the Pandemic" Washington, D.C. : World Bank Group (2023a).
- Brunckhorst, Ben, Alexandru Cojocaru, and Yeon Soo Kim. "Tracing pandemic impacts in the absence of regular survey data: What have we learned from the World Bank's high-frequency phone surveys? Washington, D.C.: World Bank Group (2023b).
- Burke, Marshall, and David B. Lobell. "Satellite-based assessment of yield variation and its determinants in smallholder African systems." *Proceedings of the National Academy of Sciences* 114, no. 9 (2017): 2189-2194.
- Burke, Marshall, Anne Driscoll, David B. Lobell, and Stefano Ermon. "Using satellite imagery to understand and promote sustainable development." *Science* 371, no. 6535 (2021): eabe8628.
- Chen, Xi, and William D. Nordhaus. "Using luminosity data as a proxy for economic statistics." *Proceedings of the National Academy of Sciences* 108, no. 21 (2011): 8589-8594.
- Caruso, German Daniel, Leonardo Lucchetti, Eduardo A. Malasquez, Thiago Scot, and Raul Castaneda. "But... what is the poverty rate today? testing poverty nowcasting methods in Latin America and the Caribbean." *World Bank Policy Research Paper* (2017).
- Chen, X. and Nordhaus, W.D., 2019. VIIRS nighttime lights in the estimation of cross-sectional and time-series GDP. *Remote Sensing*, 11(9), p.1057.

- Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., ... & Wernerfelt, N. "Social capital I: measurement and associations with economic mobility." *Nature*, 608(7921), 108-121 (2022a).
- Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., ... & Wernerfelt, N. "Social capital II: determinants of economic connectedness." *Nature*, 608(7921), 122-134 (2022b).
- Chi, Guanghua, Han Fang, Sourav Chatterjee, and Joshua E. Blumenstock. "Microestimates of wealth for all low-and middle-income countries." *Proceedings of the National Academy of Sciences* 119, no. 3 (2022): e2113658119.
- Christiaensen, Luc, Peter Lanjouw, Jill Luoto, and David Stifel. "Small area estimation-based prediction methods to track poverty: validation and applications." *The Journal of Economic Inequality* 10 (2012): 267-297.
- Cojocar, Alexandru, Uche Ekhaton-Mobayode. "Rising food and energy prices and their welfare implications: a resource note" Washington, DC: World Bank (2022)
- Cooper, Matthew W., Molly E. Brown, Stefan Hochrainer-Stigler, Georg Pflug, Ian McCallum, Steffen Fritz, Julie Silva, and Alexander Zvoleff. "Mapping the effects of drought on child stunting." *Proceedings of the National Academy of Sciences* 116, no. 35 (2019): 17219-17224.
- Corral, Paul, Isabel Molina, Alexandru Cojocar, Sandra Segovia. "Guidelines to Small Area Estimation for Poverty Mapping." Washington, DC: World Bank (2022)
- Corral, Paul, Heath Henderson, and Sandra Segovia. "Poverty Mapping in the Age of Machine Learning." Available at SSRN 4587156 (2023).
- Corral, Paul. "Should You Impute That? A brief look into the art of poverty imputation." (*forthcoming*).
- Cuesta, Jose, and Cristian Chagalj. "Measuring poverty with administrative data in data deprived contexts: The case of Nicaragua." *Economics Letters* 183 (2019): 108573.

- Cuesta, Jose, and Gabriel Lara Ibarra. "Comparing cross-survey micro imputation and macro projection techniques: Poverty in post revolution Tunisia." *Journal of Income Distribution* 25, no. 1 (2017): 1-30.
- D'Amuri, Francesco, and Juri Marcucci. "The predictive power of Google searches in forecasting US unemployment." *International Journal of Forecasting* 33, no. 4 (2017): 801-816.
- Dampha, Nfamara K., Colette Salemi, and Stephen Polasky. "Rohingya Refugee Camps and Forest Loss in Cox's Bazar, Bangladesh." (2022).
- Dang, Hai-Anh, and Minh Nguyen. "POVIMP: stata module to provide poverty estimates in the absence of actual consumption data." (2014).
- Dang, Hai-Anh H., Peter F. Lanjouw, and Umar Serajuddin. "Updating poverty estimates in the absence of regular and comparable consumption data: methods and illustration with reference to a middle-income country." *Oxford Economic Papers* 69, no. 4 (2017): 939-962.
- Dang, Hai-Anh, Dean Jolliffe, and Calogero Carletto. "Data gaps, data incompatibility, and data imputation: A review of poverty measurement methods for data-scarce environments." *Journal of Economic Surveys* 33, no. 3 (2019): 757-797.
- Dang, Hai-Anh, and Paolo Verme. "Estimating Poverty for Refugee Populations: Can Cross-Survey Imputation Methods Substitute for Data Scarcity?" *World Bank Policy Research Working Paper* 9076 (2019).
- Dang, Hai-Anh H. "To impute or not to impute, and how? A review of poverty-estimation methods in the absence of consumption data." *Development Policy Review* 39, no. 6 (2021): 1008-1030.
- Dang, Hai-Anh, Talip Kilic, Kseniya Abanokova, and Calogero Carletto. "Poverty imputation in contexts without consumption data: a revisit with further refinements." (2023).
- Debbich, M., 2019. *Assessing Oil and Non-Oil GDP Growth from Space: An Application to Yemen 2012-17*. International Monetary Fund.

- Decuyper, Adeline, Alex Rutherford, Amit Wadhwa, Jean-Martin Bauer, Gautier Krings, Thoralf Gutierrez, Vincent D. Blondel, and Miguel A. Luengo-Oroz. “Estimating food consumption and poverty indices with mobile phone data.” arXiv preprint arXiv:1412.2595 (2014).
- Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. “What do we learn from the weather? The new climate-economy literature.” *Journal of Economic literature* 52, no. 3 (2014): 740-798.
- Do, Quy-Toan, Jacob N. Shapiro, Christopher D. Elvidge, Mohamed Abdel-Jelil, Daniel P. Ahn, Kimberly Baugh, Jamie Hansen-Lewis, Mikhail Zhizhin, and Morgan D. Bazilian. “Terrorism, geopolitics, and oil security: Using remote sensing to estimate oil production of the Islamic State.” *Energy research & social science* 44 (2018): 411-418.
- Doan, Miki Khanh, Ruth Hill, Stephane Hallegatte, Paul Corral, Ben Brunckhorst, Minh Nguyen, Samuel Freije-Rodriguez, and Esther Naikal. “Counting People Exposed to, Vulnerable to, or at High Risk From Climate Shocks: A Methodology.” *World Bank Policy Research Papers* (2023).
- Doherty, Grace Anna, Andrii Vasilevich Berdnyk, Gabriel Gene Levin. “Remote Sensing: Guide to Practitioners.” Washington, D.C.: World Bank Group (2022)
- Dollar, David, and Aart Kraay. “Growth is Good for the Poor.” *Journal of economic growth* 7 (2002): 195-225.
- Doudich, Mohamed, Abdeljaouad Ezzrari, Roy Van der Weide, and Paolo Verme. “Estimating quarterly poverty rates using labor force surveys: a primer.” *The World Bank Economic Review* 30, no. 3 (2016): 475-500.
- Dubois, Pierre, Rachel Griffith, and Martin O’Connell. “The use of scanner data for economics research.” *Annual Review of Economics* 14 (2022): 723-745.
- Edochie, Ifeanyi Nzegwu, Samuel Freije-Rodriguez, Christoph Lakner, Laura Moreno Herrera, David Locke Newhouse, Sutirtha Sinha Roy, and Nishant Yonzan. “What do we Know about Poverty in India in 2017/18?” *World Bank Policy Research Paper* 9931 (2022).
- Elbers, Chris, Jean O. Lanjouw, and Peter Lanjouw. “Micro-level estimation of poverty and inequality.” *Econometrica* 71, no. 1 (2003): 355-364.

- Elvidge, Christopher D., Paul C. Sutton, Tilottama Ghosh, Benjamin T. Tuttle, Kimberly E. Baugh, Budhendra Bhaduri, and Edward Bright. "A global poverty map derived from satellite data." *Computers & Geosciences* 35, no. 8 (2009): 1652-1660.
- Elvidge, C.D., Baugh, K., Zhizhin, M., Hsu, F.C. and Ghosh, T., 2017. VIIRS night-time lights. *International journal of remote sensing*, 38(21), pp.5860-5879.
- Engstrom, Ryan, Jonathan Hersh, and David Newhouse. "Poverty from space: Using high resolution satellite imagery for estimating economic well-being." *The World Bank Economic Review* 36, no. 2 (2022): 382-412.
- Essama-Nssah, Boniface. "The poverty and distributional impact of macroeconomic shocks and policies: A review of modeling approaches." (2005).
- Fatehkia, Masoomali, Isabelle Tingzon, Ardie Orden, Stephanie Sy, Vedran Sekara, Manuel Garcia-Herranz, and Ingmar Weber. "Mapping socioeconomic indicators using social media advertising data." *EPJ Data Science* 9, no. 1 (2020): 22.
- Fezzi, Carlo, and Valeria Fanghella. "Tracking GDP in real-time using electricity market data: Insights from the first wave of COVID-19 across Europe." *European economic review* 139 (2021): 103907.
- Fraiberger, Samuel P., Pablo Astudillo, Lorenzo Candeago, Alex ChUNET, Nicholas KW Jones, Maham Faisal Khan, Bruno Lepri et al. "Uncovering socioeconomic gaps in mobility reduction during the COVID-19 pandemic using location data." *arXiv preprint arXiv:2006.15195* (2020).
- Figari, Francesco, Alari Paulus, and Holly Sutherland. "Microsimulation and policy analysis." In *Handbook of income distribution*, vol. 2, pp. 2141-2221. Elsevier, 2015.
- Finn, Arden, and Vimal Ranchhod. "Genuine fakes: The prevalence and implications of data fabrication in a large South African survey." *The World Bank Economic Review* 31, no. 1 (2017): 129-157.
- Fotheringham, A. Stewart, and David WS Wong. "The modifiable areal unit problem in multivariate statistical analysis." *Environment and planning A* 23, no. 7 (1991): 1025-1044.

- Fujii, Tomoki, and Roy van der Weide. "Is predicted data a viable alternative to real data?" *The World Bank Economic Review* 34, no. 2 (2020): 485-508.
- Gascoigne, Jon, Sandra Baquie, Katja Vinha, Emmanuel Skoufias, Evie Calcutt, Varun Kshirsagar, Conor Meenan, Ruth Hill. "The Welfare Cost of Drought in Sub-Saharan Africa". The World Bank (forthcoming)
- Gao, Jia, and Gabriela Inchauste. "A Customizable Microsimulation Tool to Analyze Distributional Effects of Country Fiscal Policies." (2020).
- Gao, Jia, Katja Vinha, and Emmanuel Skoufias. "World Bank Equity Policy Lab Vulnerability Tool to Measure Poverty Risk." (2020).
- Giannone, Domenico, Lucrezia Reichlin, and David Small. "Nowcasting: The real-time informational content of macroeconomic data." *Journal of monetary economics* 55, no. 4 (2008): 665-676.
- Gibson, John, Susan Olivia, and Geua Boe-Gibson. "Night lights in economics: Sources and uses 1." *Journal of Economic Surveys* 34, no. 5 (2020): 955-980.
- Gibson, John, Susan Olivia, Geua Boe-Gibson, and Chao Li. "Which night lights data should we use in economics, and where?" *Journal of Development Economics* 149 (2021): 102602.
- Gourlay, Sydney, Talip Kilic, Antonio Martuscelli, Philip Wollburg, and Alberto Zezza. "High-frequency phone surveys on COVID-19: good practices, open questions." *Food Policy* 105 (2021): 102153.
- Gualavisi, Melany and David Locke Newhouse. "Integrating Survey and Geospatial Data to Identify the Poor and Vulnerable: Evidence from Malawi." Washington, D.C.: World Bank Group (2022)
- Hall, Ola, Francis Dompae, Ibrahim Wahab, and Fred Mawunyo Dzanku. "A review of machine learning and satellite imagery for poverty prediction: Implications for development research and applications." *Journal of International Development* (2023).

- Habib, Bilal, Ambar Narayan, Sergio Olivieri, and Carolina Sanchez-Paramo. "Assessing ex ante the poverty and distributional impact of the global crisis in a developing country: A micro-simulation approach with application to Bangladesh." World Bank Policy Research Working Paper 5238 (2010).
- Hallegatte, Stephane, Adrien Vogt-Schilb, Mook Bangalore, and Julie Rozenberg. *Unbreakable: building the resilience of the poor in the face of natural disasters*. World Bank Publications, 2016.
- Henderson, J. Vernon, Adam Storeygard, and David N. Weil. "Measuring economic growth from outer space." *American economic review* 102, no. 2 (2012): 994-1028.
- Hentschel, Jesko. "Combining census and survey data to study spatial dimensions of poverty: a case study of Ecuador." Vol. 1928. World Bank Publications, 1998.
- Hernandez, Marco, Lingzi Hong, Vanessa Frias-Martinez, Andrew Whitby, and Enrique Frias-Martinez. "Estimating poverty using cell phone data: evidence from Guatemala." World Bank Policy Research Working Paper 7969 (2017).
- Hill, Ruth Vargas, and Catherine Porter. "Vulnerability to drought and food price shocks: evidence from Ethiopia." *World Development* 96 (2017): 65-77.
- Himelein, Kristen., Eckman, S., Kastelic, J., McGee, K., Wild, M., Yoshida, N., Hoogeveen, J. "High frequency mobile phone surveys of households to assess the impacts of COVID-19: guidelines on sampling design." Washington, D.C.: World Bank Group (2020)
- Hoogeveen, Johannes, and Utz Pape. *Data collection in fragile states: innovations from Africa and beyond*. Springer Nature, 2020.
- Hoy, Christopher; Kim, Yeon Soo; Nguyen, Minh; Sosa, Mariano; Tiwari, Sailesh. "Building Public Support for Reducing Fossil Fuel Subsidies: Evidence across 12 Middle-Income Countries. World Bank Policy Research Working Papers; 10615 (2023)

- Huang, Luna Yue, Solomon M. Hsiang, and Marco Gonzalez-Navarro. “Using satellite imagery and deep learning to evaluate the impact of anti-poverty programs.” No. w29105. National Bureau of Economic Research, 2021.
- Jaravel, Xavier, and Martin O’Connell. “Real-time price indices: Inflation spike and falling product variety during the Great Lockdown.” *Journal of Public Economics* 191 (2020): 104270.
- International Telecommunication Union (ITU). “Facts and Figures 2022” (2022)
- Jean, Neal, Marshall Burke, Michael Xie, W. Matthew Davis, David B. Lobell, and Stefano Ermon. “Combining satellite imagery and machine learning to predict poverty.” *Science* 353, no. 6301 (2016): 790-794.
- Jones, Nicholas, Takahiro Yabe, Samuel Heroy. “Cities on the Move. A Technical Guidance Note.” World Bank, Washington DC (forthcoming)
- Kakwani, Nanak. “Poverty and economic growth with application to Côte d’Ivoire.” *Review of Income and Wealth* 39, No. 2 (1993): 121-139.
- Keola, Souknilanh, Magnus Andersson, and Ola Hall. “Monitoring economic development from space: using nighttime light and land cover data to measure economic growth.” *World Development* 66 (2015): 322-334.
- Kilic, Talip, Umar Serajuddin, Hiroki Uematsu, and Nobuo Yoshida. “Costing household surveys for monitoring progress toward ending extreme poverty and boosting shared prosperity.” World Bank Policy Research Working Paper 7951 (2017).
- Kim, Yeon Soo, Jeffrey Tanner. “Displaced During Crisis. Lessons Learned from High-Frequency Phone Surveys and How to Protect the Most Vulnerable.” Washington, D.C.: World Bank Group (2023).
- Klasen, Stephan, and Mark Misselhorn. Determinants of the growth semi-elasticity of poverty reduction. No. 176. IAI Discussion Papers, 2008.
- Knippenberg, Erwin, Nathaniel Jensen, and Mark Conostas. “Quantifying household resilience with high frequency data: Temporal dynamics and methodological options.” *World Development* 121 (2019): 1-15.

- Kocornik-Mina, Adriana, Thomas KJ McDermott, Guy Michaels, and Ferdinand Rauch. "Flooded cities." *American Economic Journal: Applied Economics* 12, no. 2 (2020): 35-66.
- Kraay, Aart. "When is growth pro-poor? Evidence from a panel of countries." *Journal of development economics* 80, no. 1 (2006): 198-227.
- Kreindler, Gabriel E., and Yuhei Miyauchi. "Measuring commuting and economic activity inside cities with cell phone records." *Review of Economics and Statistics* 105, no. 4 (2023): 899-909.
- Kshirsagar, Varun, Jerzy Wieczorek, Sharada Ramanathan, and Rachel Wells. "Household poverty classification in data-scarce environments: A machine learning approach." arXiv preprint arXiv:1711.06813 (2017).
- Kugler, Maurice, Viollaz, Mariana, Duque, Daniel, Gaddis, Isis, Newhouse, David, Palacios-Lopez, Amparo, and Weber, Michael. "How Did the COVID-19 Crisis Affect Different Types of Workers in the Developing World?" Policy Research Working Paper 9703. World Bank, Washington DC (2021).
- Lain, Jonathan, Marta Schoch, and Tara Vishwanath. "Making Data Count: Estimating a Poverty Trend for Nigeria between 2009 and 2019." *The World Bank Economic Review* (2023): lhad032.
- Lakner, Christoph, Mario Negre, Espen Beer Prydz, and Mario Negre Rossignoli. "Twinning the goals: how can promoting shared prosperity help to reduce global poverty?" World Bank Policy Research Working Paper 7106 (2014).
- Lakner, Christoph, Daniel Gerszon Mahler, Mario Negre, and Espen Beer Prydz. "How much does reducing inequality matter for global poverty?" *The Journal of Economic Inequality* 20, no. 3 (2022): 559-585.
- Larrimore, Jeff, Jacob Mortenson, and David Splinter. "Earnings shocks and stabilization during COVID-19." *Journal of Public Economics* 206 (2022): 104597.
- Lazer, David, Ryan Kennedy, Gary King, and Alessandro Vespignani. "The parable of Google Flu: traps in big data analysis." *Science* 343, no. 6176 (2014): 1203-1205.
- Lee, Kamwoo, and Jeanine Braithwaite. "High-resolution poverty maps in Sub-Saharan Africa." *World Development* 159 (2022): 106028.

- Lee, Sunghee. “Propensity score adjustment as a weighting scheme for volunteer panel web surveys.” *Journal of official statistics* 22, no. 2 (2006): 329.
- Lentz, Erin C., Hope Michelson, Katherine Baylis, and Yang Zhou. “A data-driven approach improves food insecurity crisis prediction.” *World Development* 122 (2019): 399-409.
- Little, Roderick JA. “Survey nonresponse adjustments for estimates of means.” *International Statistical Review* (1986): 139-157.
- Loayza, Norman V., and Claudio Raddatz. “The composition of growth matters for poverty alleviation.” *Journal of Development Economics* 93, no. 1 (2010): 137-151.
- Mahler, Daniel Gerszon, R. Andrés Castañeda Aguilar, and David Newhouse. “Nowcasting global poverty.” *The World Bank Economic Review* 36, no. 4 (2022a): 835-856.
- Mahler, Daniel Gerszon, Nishant Yonzan, and Christoph Lakner. “The impact of COVID-19 on global inequality and poverty.” *World Bank Policy Research Working Paper* 10198 (2022b).
- Martinez, Luis R. “How much should we trust the dictator’s GDP growth estimates?” *Journal of Political Economy* 130, no. 10 (2022): 2731-2769.
- Marty, Robert, and Alice Duhaut. “Global poverty estimation using private and public sector big data sources.” *Scientific Reports* 14, no. 1 (2024): 3160.
- Mathiassen, A., 2009. A model-based approach for predicting annual poverty rates without expenditure data. *The Journal of Economic Inequality*, 7, pp.117-135.
- McBride, Linden, Christopher B. Barrett, Christopher Browne, Leiqiu Hu, Yanyan Liu, David S. Matteson, Ying Sun, and Jiaming Wen. “Predicting poverty and malnutrition for targeting, mapping, monitoring, and early warning.” *Applied Economic Perspectives and Policy* 44, no. 2 (2022): 879-892.
- Merfeld, Joshua D., and David Newhouse. “Improving Estimates of Mean Welfare and Uncertainty in Developing Countries.” No. 10348. *The World Bank*, 2023.

- Michalopoulos, Stelios, and Elias Papaioannou. "Spatial patterns of development: A meso approach." *Annual Review of Economics* 10 (2018): 383-410.
- Milusheva, Sveta, Robert Marty, Guadalupe Bedoya, Sarah Williams, Elizabeth Resor, and Arianna Legovini. "Applying machine learning and geolocation techniques to social media data (Twitter) to develop a resource for urban planning." *PloS one* 16, no. 2 (2021): e0244317.
- Minetto, Rodrigo, Mauricio Pamplona Segundo, Gilbert Rotich, and Sudeep Sarkar. "Measuring human and economic activity from satellite imagery to support city-scale decision-making during COVID-19 pandemic." *IEEE Transactions on Big Data* 7, no. 1 (2020): 56-68.
- Mohammed, Gina H., Roberto Colombo, Elizabeth M. Middleton, Uwe Rascher, Christiaan van der Tol, Ladislav Nedbal, Yves Goulas et al. "Remote sensing of solar-induced chlorophyll fluorescence (SIF) in vegetation: 50 years of progress." *Remote sensing of environment* 231 (2019): 111177.
- Montoya Munoz, Kelly Yelitza, Sergio Daniel Olivieri, and Cicero Augusto Silveira Braga. "Considering Labor Informality in Forecasting Poverty and Inequality: A Microsimulation Model for Latin American and Caribbean Countries." No. 10497. The World Bank, 2023.
- Morgandi, Matteo, Katharina Maria Fietz, Malin Linnea Sofia Ed, and Gabriel Fagundes De Oliveira. "From Struggle to Opportunity-The Profile of Brazil's Working Poor and Implications for Economic Inclusion." (2023).
- Newhouse, David Locke, Shivapragasam Shivakumaran, Shinya Takamatsu, and Nobuo Yoshida. "How survey-to-survey imputation can fail." *World Bank Policy Research Working Paper* 6961 (2014).
- Newhouse, David. "Small Area Estimation of Poverty and Wealth Using Geospatial Data: What Have We Learned So Far?" (2023).
- Nguyen, Minh, Paul Andres Corral Rodas, João Pedro Azevedo, and Qinghua Zhao. "sae: A stata package for unit level small area estimation." *World Bank Policy Research Working Paper* 8630 (2018).

- Olivieri, Sergio, Sergiy Radyakin, Stanislav Kolenikov, Michael Lokshin, Ambar Narayan, and Carolina Sanchez-Paramo. "Simulating distributional impacts of macro-dynamics: theory and practical applications." World Bank Publications, 2014.
- Olivieri, Sergio, Francesc Ortega, Ana Rivadeneira, and Eliana Carranza. "The labour market effects of Venezuelan migration in Ecuador." *The Journal of Development Studies* 58, no. 4 (2022): 713-729.
- Palotti, Joao, Natalia Adler, Alfredo Morales-Guzman, Jeffrey Villaveces, Vedran Sekara, Manuel Garcia Herranz, Musa Al-Asad, and Ingmar Weber. "Monitoring of the Venezuelan exodus through Facebook's advertising platform." *PLOS ONE*, 15.2 (2020): e0229175.
- Pape, Utz Johann, and Johan A. Mistiaen. "Household expenditure and poverty measures in 60 minutes: a new approach with results from Mogadishu." World Bank Policy Research Working Paper 8430 (2018).
- Pape, Utz Johann. "Measuring poverty rapidly using within-survey imputations." Policy Research Working Papers. World Bank (2021).
- Picot, Garnett, and Patrizio Piraino. "Immigrant earnings growth: selection bias or real progress?" *Canadian Journal of Economics/Revue canadienne d'économique* 46, no. 4 (2013): 1510-1536.
- Pieterse, Duncan, Elizabeth Gavin, and C. Friedrich Kreuser. "Introduction to the South African Revenue Service and National Treasury Firm-Level Panel." *South African Journal of Economics* 86 (2018): 6-39.
- Pinkovskiy, Maxim, and Xavier Sala-i-Martin. "Lights, camera... income! Illuminating the national accounts-household surveys debate." *The Quarterly Journal of Economics* 131, no. 2 (2016): 579-631.
- Pokhriyal, Neeti, and Damien Christophe Jacques. "Combining disparate data sources for improved poverty prediction and mapping." *Proceedings of the National Academy of Sciences* 114, no. 46 (2017): E9783-E9792.
- Pople, Ashley, Ruth Hill, Stephan Dercon, and Ben Brunckhorst. "Anticipatory Cash Transfers in Climate Disaster Response." CSAE Working Paper Series, Centre for the Study of African Economies (2021).

- Proctor, Jonathan, Tamma Carleton, and Sandy Sum. "Parameter recovery using remotely-sensed variables." No. w30861. National Bureau of Economic Research (2023).
- Prydz, Espen Beer, Dean Jolliffe, and Umar Serajuddin. "Disparities in Assessments of Living Standards Using National Accounts and Household Surveys." *Review of Income and Wealth* 68 (2022): S385-S420.
- Ratledge, Nathan, Gabe Cadamuro, Brandon de la Cuesta, Matthieu Stigler, and Marshall Burke. «Using machine learning to assess the livelihood impact of electricity access.» *Nature* 611, no. 7936 (2022): 491-495.
- Ravallion, Martin. "Growth and poverty: Evidence for developing countries in the 1980s." *Economics letters* 48, no. 3-4 (1995): 411-417.
- Ravallion, Martin, and Shaohua Chen. "What can new survey data tell us about recent changes in distribution and poverty?" *The World Bank Economic Review* 11, no. 2 (1997): 357-382.
- Ravallion, Martin. "Poverty lines in theory and practice." Vol. 133. World Bank Publications, 1998.
- Ravallion, Martin. "Growth, inequality and poverty: looking beyond averages." *World development* 29, no. 11 (2001): 1803-1815.
- Ravallion, Martin. "Measuring aggregate welfare in developing countries: How well do national accounts and surveys agree?" *Review of Economics and Statistics* 85, no. 3 (2003): 645-652.
- Rivas, Lisbeth, and Joe Crowley. "Using Administrative Data to Enhance Policymaking in Developing Countries: Tax Data and the National Accounts." International Monetary Fund, 2018.
- Rivera, Nayib, Mehtabul Azam, and Mohamed Ihsan Ajwad. "Tracing Labor Market Outcomes of Technical and Vocational Training Graduates in Saudi Arabia." (2022).
- Rodríguez-Castelán, Carlos, Abdelkrim Araar, Eduardo A. Malásquez, Sergio Daniel Olivieri, and Tara Vishwanath. "Distributional effects of competition: A simulation approach." World Bank Policy Research Working Paper 8838 (2019).

- Rolf, Esther, Jonathan Proctor, Tamma Carleton, Ian Bolliger, Vaishaal Shankar, Miyabi Ishihara, Benjamin Recht, and Solomon Hsiang. “A generalizable and accessible approach to machine learning with global satellite imagery.” *Nature communications* 12, no. 1 (2021): 4392.
- Rolf, Esther. “Evaluation Challenges for Geospatial ML.” arXiv preprint arXiv:2303.18087 (2023).
- McKenzie, David, and Dario Sansone. “Predicting entrepreneurial success is hard: Evidence from a business plan competition in Nigeria.” *Journal of Development Economics* 141 (2019): 102369.
- Roson, Roberto, and Martina Sartori. „Estimation of climate change damage functions for 140 regions in the GTAP9 database.“ *Journal of Global Economic Analysis* (2016).
- Sinha Roy, Sutirtha, and Roy Van Der Weide. “Poverty in India Has Declined over the Last Decade But Not As Much As Previously Thought.” *World Bank Policy Research Paper* (2022).
- Skoufias, Emmanuel, Alexis Diamond, Katja Vinha, Michael Gill, and Miguel Rebolledo Dellepiane. “Estimating poverty rates in subnational populations of interest: An assessment of the Simple Poverty Scorecard.” *World Development* 129 (2020): 104887.
- Skoufias, Emmanuel, Eric Strobl, and Thomas Tveit. “Can we rely on VIIRS night-lights to estimate the short-term impacts of natural hazards? Evidence from five South East Asian countries.” *Geomatics, natural hazards and risk* 12, no. 1 (2021): 381-404.
- Smythe, Isabella S., and Joshua E. Blumenstock. “Geographic microtargeting of social assistance with high-resolution poverty maps.” *Proceedings of the National Academy of Sciences* 119, no. 32 (2022): e2120025119.
- Steele, Jessica E., Pål Roe Sundsøy, Carla Pezzulo, Victor A. Alegana, Tomas J. Bird, Joshua Blumenstock, Johannes Bjelland et al. “Mapping poverty using mobile phone and satellite data.” *Journal of The Royal Society Interface* 14, no. 127 (2017): 20160690.

- Stifel, David, and Luc Christiaensen. "Tracking poverty over time in the absence of comparable consumption data." *The World Bank Economic Review* 21, no. 2 (2007): 317-341.
- Tabakis, Chrysostomos, Gi Khan Ten, Joshua D. Merfeld, David Newhouse, Utz Pape, and Michael Weber. "The Welfare Implications of COVID-19 for Fragile and Conflict-Affected Areas." (2022).
- Tang, Binh, Yanyan Liu, and David S. Matteson. "Predicting poverty with vegetation index." *Applied Economic Perspectives and Policy* 44, no. 2 (2022): 930-945.
- Taptué, Andre-Marie, and Johannes Hooegeven. "Resident enumerators for continuous monitoring." *Data collection in fragile states: Innovations from Africa and beyond* (2020): 63-82.
- Tarozzi, A., 2007. Calculating comparable statistics from incomparable surveys, with an application to poverty in India. *Journal of business & economic statistics*, 25(3), pp.314-336.
- Ten, Gi Khan, Josh Merfeld, David Newhouse, Utz Pape, and Kibrom Taferes Hirrfot. "How Well Can Real-Time Indicators Track the Economic Impacts of a Crisis Like COVID-19?" (2022).
- Van Der Weide, Roy, Brian Blankespoor, Chris Elbers, and Peter Lanjouw. "How Accurate Is a Poverty Map Based on Remote Sensing Data? An Application to Malawi." *World Bank Policy Research Paper* (2023).
- Van Der Weide, Roy, Christoph Lakner, and Elena Ianchovichina. "Is inequality underestimated in Egypt? Evidence from house prices." *Review of Income and Wealth* 64 (2018): S55-S79.
- Verschuur, Jasper, Elco E. Koks, and Jim W. Hall. "Global economic impacts of COVID-19 lockdown measures stand out in high-frequency shipping data." *PloS one* 16, no. 4 (2021): e0248818.
- Viollaz, Mariana, Daniel Duque, Carolina Diaz-Bonilla, David Newhouse, and Michael Weber. "From Middle Class to Poverty: The Unequal Impacts of the COVID-19 Pandemic on Developing Countries". *Policy Research Working Papers*. World Bank (2023)

- Voukelatou, Vasiliki, Lorenzo Gabrielli, Ioanna Miliou, Stefano Cresci, Rajesh Sharma, Maurizio Tesconi, and Luca Pappalardo. “Measuring objective and subjective well-being: dimensions and data sources.” *International Journal of Data Science and Analytics* 11 (2021): 279-309.
- Walsh, Brian, and Stephane Hallegatte. “Measuring natural risks in the Philippines: socioeconomic resilience and wellbeing losses.” *Economics of Disasters and Climate Change* 4 (2020): 249-293.
- Watmough, Gary R., Charlotte LJ Marcinko, Clare Sullivan, Kevin Tschirhart, Patrick K. Mutuo, Cheryl A. Palm, and Jens-Christian Svenning. “Socioecologically informed use of remote sensing data to predict rural household poverty.” *Proceedings of the National Academy of Sciences* 116, no. 4 (2019): 1213-1218.
- World Bank Group, and UK Department of International Development. “Poverty and Vulnerability in the Ethiopian Lowlands: Building a More Resilient Future.” (2019).
- World Bank. “Stress Testing Social Protection. A guide for practitioners”. Washington, D.C.: World Bank Group (2021a)
- World Bank. “World Development Report 2021: Data for better lives”. Washington, D.C.: World Bank Group (2021b)
- Yeh, Christopher, Anthony Perez, Anne Driscoll, George Azzari, Zhongyi Tang, David Lobell, Stefano Ermon, and Marshall Burke. “Using publicly available satellite imagery and deep learning to understand economic well-being in Africa.” *Nature communications* 11, no. 1 (2020): 2583.
- Yonzan, Nishant, Branko Milanovic, Salvatore Morelli, and Janet Gornick. “Drawing a line: comparing the estimation of top incomes between tax data and household survey data.” *The Journal of Economic Inequality* 20, no. 1 (2022): 67-95.
- Yoshida, Nobuo, Hiroki Uematsu, and Carlos E. Sobrado. “Is extreme poverty going to end? An analytical framework to evaluate progress in ending extreme poverty.” *An Analytical Framework to Evaluate Progress in Ending Extreme Poverty* (January 1, 2014). World Bank Policy Research Working Paper 6740 (2014).

- Yoshida, Nobuo, R. Munoz, A. Skinner, C. Kyung-eun Lee, M. Brataj, W. Durbin, D. Sharma, and C. Wieser. "Swift data collection guidelines version 2." The World Bank (2015).
- Yoshida, Nobuo, Shinya Takamatsu, Shivapragasam Shivakumaran, Danielle Aron, Xueqi Li, and Kazusa Yoshimura. "Frequent and timely monitoring of poverty, inequality, and poverty profiles using SWIFT during the COVID-19 Pandemic." (2020)
- Yoshida, Nobuo, X. Chen, S. Takamatsu, K. Yoshimura, S. Malgioglio, and S. Shivakumaran. "The Concept and Empirical Evidence of SWIFT Methodology". World Bank (2022).
- Yoshida, Nobuo, Danielle Aron. "Enabling High-frequency and Real-time Poverty Monitoring in the Developing World with SWIFT (Survey of Wellbeing via Instant and Frequent Tracking)". World Bank (*forthcoming*).
- Yoshimura, Kazusa, Danielle Aron, James Campbell, Xueqi Li, Joanna Upton, Nobuo Yoshida, and Kexin Zhang. "Rapid Feedback Monitoring System (RFMS)—real-time, cost-effective, shock-resilient monitoring of living conditions and food security." World Bank (2022).
- Zeza, Alberto; Mcgee, Kevin; Wollburg, Philip; Assefa, Thomas; Gourlay, Sydney. "From Necessity to Opportunity: Lessons for Integrating Phone and In-Person Data Collection for Agricultural Statistics in a Post-Pandemic World." Policy Research Working Paper 10168. World Bank, Washington D (2022)
- Zhang, Qingling, and Karen C. Seto. "Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data." *Remote Sensing of Environment* 115, no. 9 (2011): 2320-2329.
- Zhang, Kexin, Shinya Takamatsu, and Nobuo Yoshida. "Correcting Sampling and Nonresponse Bias in Phone Survey Poverty Estimation Using Reweighting and Poverty Projection Models" World Bank (2023)
- Zhao, Xizhi, Bailang Yu, Yan Liu, Zuoqi Chen, Qiaoxuan Li, Congxiao Wang, and Jianping Wu. "Estimation of poverty using random forest regression with multi-source data: A case study in Bangladesh." *Remote Sensing* 11, no. 4 (2019): 375.

Annex 1.

Summary of Models Used to Update Poverty Estimates

Which model is best suited ultimately depends on a range of factors, including the specific research question, the geographic scope, availability of baseline and auxiliary data, but also the available time to implement the model. Table A1 below provides a short summary as well as limitations of the different models discussed in the first part of the typology.

TABLE A1 Summary and limits of models presented in this typology

| | Summary | Limitations |
|---|--|--|
| Survey- and non-survey covariate based nowcasting | <ul style="list-style-type: none"> • Imputation of consumption/income based on a recent auxiliary data and a baseline survey • The baseline survey needs to have a full consumption or expenditure module • The auxiliary survey dataset must have predictor variables that can be matched to the baseline survey • If no auxiliary data is available, the collection of 10-15 covariates can be sufficient to impute poverty,- • Non-survey data can also complement auxiliary data in survey-to-survey imputation. • For non-survey data, oftentimes machine learning models are used. | <ul style="list-style-type: none"> • Need for consistent measurement of predictors between baseline and auxiliary survey • Cannot be used to analyze poverty rates beyond the sampling design/ coverage of the auxiliary dataset • Cannot be used when the relationship between predictor and welfare variables is different between baseline and auxiliary data collection. • Machine-learning models can be complex and “black-boxes”, which diminishes the interpretability of obtained estimates • “Off-the-shelf” applications are rare, and models are usually very specific to the context of the training (ground-truth) data |

| | Summary | Limitations |
|--------------------------------|---|---|
| Nowcasting based on GDP growth | <ul style="list-style-type: none"> • A relatively simple method of estimating poverty rates, based on current and past GDP (or Household Final Consumption Expenditure) • The two main aretypes are distribution-scaling and the GDP-poverty elasticity method. | <ul style="list-style-type: none"> • In many settings, the linkages between growth and poverty are complex, and past elasticities are not always accurate • Not best suited to study changes in income distributions • By itself, cannot measure heterogeneous effects of shocks on different sub-populations. Though macro-decomposition tools can be used to reweight effects on different groups. |
| Microsimulation models | <ul style="list-style-type: none"> • Microsimulation models combine macroeconomic projections and relationships with microeconomic distributions • These types of models are useful to study channels of how changes and shocks affect poverty, and distributional effects • Several standalone models exist for different regions, and to investigate specific types of scenarios, such as climate shocks, inflation, or fiscal policies. • More complex macro-micro simulation models incorporate Computable General Equilibrium models • Models such as GIDD allow to incorporate indirect effects as well as a broad range of channels | <ul style="list-style-type: none"> • Macro-micro models generally simulate future changes or potential shocks, but do not provide ex-post assessments of actual shocks • Requires modelling assumptions, as well as sufficient data. |

Annex 2.

Commonly Used Machine Learning (ML) Models for Estimating Poverty

TABLE 2 ML models for poverty estimation

| Description | Example Applications | |
|---|--|--|
| Supervised Machine Learning Methods | | |
| Least Absolute Shrinkage and Selection Operator (LASSO)/ Ridge Regressions/ Elastic Nets | LASSO and ridge regressions are comparatively simple methods for model selection and parameter estimation. They regularize and select variables by penalizing features with low predictive power. LASSO and ridge regressions are most often used to select variables and features with the largest predictive power, which are then used in other ML or linear regression models. Elastic Nets combine LASSO and Ridge Regression methods to achieve both feature selection and coefficient shrinkage, and perform better when features exhibit high levels of correlation. | Knippenberg et al. (2019) use LASSO and Random Forest algorithms to filter the strongest predictors of food insecurity. Kshirsagar et al. (2017) use an elastic net approach to select ten questions that identify poor households in Zambia. Marty and Duhaut (2024) use LASSO, XGBoost and support vector machines to predict welfare. |
| Decision Trees and Random Forests (RF) | Decision trees divide samples in a hierarchical tree-like structure based on observed features, which are easy to interpret. Random forests are an ensemble method averaging multiple decision trees, which reduces overfitting and improves performance, but is less easily interpretable. Can be used for classification or regression. | Andree et al. (2020) use RF to predict food crises. Browne et al. (2021) use RF to predict malnutrition and poverty rates with survey and satellite data. Watmough et al. (2019) and Zhao et al. (2019) train RF models on spatial data to predict poverty rates in Kenya, Bangladesh, and Zambia. |

| | Description | Example Applications |
|--|--|--|
| Gradient Boosting | Gradient boosting is, similar to random forests, an ensemble method that combines multiple weak predictors (for example, decision trees). Boosting creates strong predictive models (typically stronger than random forest models), by iteratively including residuals of previous iterations in predictions. The results are not easily interpretable. Can be used for classification or regression. | Corral et al. (2023) show that gradient boosting compares to traditional poverty mapping methods only when the underlying data is of high quality. Chi et al. (2022) use boosting methods to develop a wealth measure for 135 low- and middle-income countries. Aiken et al. (2022) test ML methods for the targeting of social assistance. |
| Support Vector Machines (SVM) and Kernels | SVM algorithms maps observations in a space with as many dimensions as there are features in the data through kernel functions. The algorithm classifies observations into two groups, by identifying a hyperplane, which maximizes the distance between these two groups. SVM algorithms are useful for text classification problems (see for example, Milusheva et al. (2021)). ¹ | Milusheva et al. (2021) use SVM to identify posts on Twitter that capture locations of crashes. McKenzie and Sansone (2019) evaluate different ML and regression models in the context of a business plan competition in Nigeria. |
| Deep Learning and Convolutional Neural Networks (CNN) | Deep Learning models extract patterns and features from large sets of input data. Deep learning models consist of multiple layers, which are connected (similar to neurons in the brain). Layers between the input and output layers are called hidden layers, which process the data with complex non-linear models. CNNs are a specific type of deep learning model that are particularly useful for image classification and therefore for applications with satellite data. Several CNN models have been developed, which can be openly accessed (see cited literature). In most cases, models are trained with labeled data, though applications in unsupervised machine learning contexts also exist. Time to train models can be long for large datasets. | Jean et al. (2016) and Yeh et al. (2020) use deep learning methods on satellite imagery and household surveys to predict welfare in local areas in African countries. Huang et al. (2021) evaluate an anti-poverty program in Kenya and Ratledge et al. (2022) apply CNN on survey and satellite data, finding that electrification increases asset wealth in rural Uganda. Tang et al. (2022) show that CNN trained on vegetation indices can predict poverty in real time in agricultural societies. |

¹ Naïve Bayes is another type of algorithm for text classification, which is easier to interpret, but typically does not perform better than SVM.

| | Description | Example Applications |
|--|---|--|
| Mixed Models and Transfer Learning | Different types of ML models can be mixed to improve model performance. For example, Lee and Braithwaite (2022) use an iterative process of predicting wealth classes using extreme gradient boosting based on labeled spatial features and using a CNN model with unlabeled satellite imageries. Transfer Learning is a technique that uses a trained model as an input in another model. Such a process can be useful when there is no labeled training data. The first model creates features that are then fed into the second model. | Jean et al. (2016) and Tang et al. (2022) circumvent the lack of labeled data by training a CNN on nighttime lights and daytime satellite data (or vegetation indices), to create intermediate labels. These labels are fed into a second machine learning model to predict wealth indicators. |
| Gaussian Process Regressions (GPR) | GPR is a non-parametric supervised machine learning method for regression and classification tasks. The probabilistic, nonlinear model yields point as well as uncertainty estimates. These uncertainty estimates may be useful for traditional estimation procedures for poverty and inequality rates. There are only a few applications in Economics. | Pokhriyal and Jacques (2017) apply a GPR model to predict multidimensional poverty indicators, using both environmental and call records data. Models are selected through Elastic Net procedures, to prevent overfitting. |
| Unsupervised Machine Learning Methods | | |
| K-Means Clustering | K-Means Clustering is a simple unsupervised method, which clusters data into k groups, based on similarity of observations across different variables by minimizing the squared Euclidean distance. | Abdul Rahman et al. (2021) use a K-Means Clustering approach to create different indicators of Multidimensional Poverty. |
| Principal Component Analysis (PCA) | PCA is a commonly used, unsupervised technique to reduce dimensionality in the data by creating features with the largest variance. Created components are orthogonal to each other. PCA is often applied for the construction of wealth indices, by combining a large set of variables to fewer components. | Ratlidge et al. (2022) use PCA to create an asset wealth index based on survey responses to 12 asset-related questions. This index is used as an outcome in welfare prediction. |

Annex 3.

Summary of All Data Sources

To reduce both time and cost of welfare estimation, practitioners and researchers have tapped alternative data sources, which allow for a greater frequency of data collection, or to obtain poverty estimates in data-deprived contexts. High-frequency data collection methods, especially the implementation of phone surveys, have proliferated during the COVID-19 pandemic, when traditional data collection methods were not feasible. The advancement of new methods and computational power (see part 1 of this typology) is also increasing the feasibility to tap large, complex, and diverse data sources, typically beyond the capacity of traditional data processing and management tools. Each dataset comes with its own advantages and caveats. Table 3 summarizes the main data sources collected and utilized by the Poverty and Equity GP to support real-time welfare monitoring efforts. Special emphasis is placed on the suitability for generating accurate and representative poverty and welfare estimates, as well as the key considerations that practitioners and researchers need to bear in mind when utilizing these data sources.

TABLE 3 Data sources for real-time welfare monitoring, advantages and disadvantages

| High Frequency Data Collection | | Harnessing Existing Data Sources | | | |
|--------------------------------|---|---|---|--|---|
| Phone surveys | Face-to-Face and online/messaging-based surveys | Geospatial data | Digital trace data | Administrative data | |
| Advantages | <ul style="list-style-type: none"> • Data collection is faster and cheaper than traditional collection. This makes phone surveys particularly useful when: • Data is needed at high frequency, and not all variables from a full household survey are required. • In contexts where in-person data collection is difficult, for example, due to security or health concerns. • When data is needed urgently, without much delay between survey planning, collection, and obtaining final data. | <ul style="list-style-type: none"> • Similar to phone surveys, high-frequency face-to-face surveys are useful when data is needed urgently or repeatedly, and full household surveys are not needed. • Community-driven or other types of face-to-face surveys overcome some of the caveats related to phone surveys, related to sampling, and to non-response bias. • Other types of surveys, for instance, implemented online or via SMS, can generate rapid information in fast-changing environments. | <ul style="list-style-type: none"> • High granularity and global coverage make geospatial data valuable in data-deprived contexts. • There are many different types of geospatial data, that can be used to proxy different economic activities. • Data is often readily available and can be linked to survey data or other data sources. | <ul style="list-style-type: none"> • Phone record data contains vast amounts of information, at high frequency and across large populations. • It can complete survey data by providing additional otherwise unmeasured aspects. • Compared to geospatial data, phone and social media data often includes information about individuals' behavior directly. • Certain aspects, such as mobility, can be directly tracked using these forms of big data. | <ul style="list-style-type: none"> • Administrative data can represent accurate depictions of income and wealth of populations, such as tax statements. • Administrative data oftentimes exists, but is seldom tapped. • Specific types of data are recorded "automatically", such as transaction records. This enables high-frequency analyses over longer time periods. |

| High Frequency Data Collection | | Harnessing Existing Data Sources | | | |
|--------------------------------|---|--|--|--|---|
| Phone surveys | Face-to-Face and online/messaging-based surveys | Geospatial data | Digital trace data | Administrative data | |
| Disadvantages | <ul style="list-style-type: none"> • Phone surveys need to be short and focused, and often cannot incorporate full consumption modules. This can make poverty estimation difficult, without appropriate estimation methods. • There can be concerns with regards to how representative phone survey data is, and sampling can be difficult when no recent survey data exists. • Non-response bias is larger than in face-to-face surveys. | <ul style="list-style-type: none"> • For there to be a time and cost-advantage over full-fledged household surveys, high-frequency face-to-face surveys need to be shorter and cannot offer full data collection. • With reduced consumption modules, inaccuracies can arise from applying imputation methods. • Online or SMS surveys are often not representative and need to be very short. | <ul style="list-style-type: none"> • Poverty and welfare estimations based on geospatial data are not always accurate, especially when measuring changes over time rather than in a cross-section. • Translating vast amounts of data into economically meaningful metrics often requires complex machine learning models and high computational power. • Baseline survey data is required, to train models and identify relevant images. • Models that work in one place cannot always be used for other locations. • Accessing the exact location of households can violate privacy concerns. | <ul style="list-style-type: none"> • Data can be heavily skewed towards wealthier households, diminishing its usefulness for poverty monitoring. • Direct measurement of poverty and welfare has not yet proven to be accurate. For poverty and welfare measurement, its main value added arises from complementing other survey data, rather than as a standalone product. • It requires relatively recent survey data to train models, which, as with geospatial data, can be complex and difficult to disentangle. • Utilizing private data such as phone records can violate privacy concerns. | <ul style="list-style-type: none"> • That administrative data is not often used in low-income country contexts is well reasoned. Collection and storage of such data is not advanced, especially in low-income contexts. • Available data may reflect only certain population groups, such as those in the labor force, and neglect others that are particularly important for poverty estimation. • Proprietary data is often hard to come by and usage can violate privacy concerns. |

Annex 4.

Nowcasting Impacts of Shocks *(Vulnerability and Damage Functions)*

A common use case for real-time monitoring is to assess the impact of shocks, such as natural disasters, on poverty and welfare. With rising temperatures and increasingly occurring natural disasters, there is an growing interest in estimating how such shocks affect poverty rates. Such estimations are key to inform policies on recovery, such as cash transfers. A reminder is warranted here that, in this document, we focus on nowcasting and real-time monitoring, which is a different type of analysis than is used in a large part of the climate-damage literature. First, many climate-damage models aim to forecast economic damages from climate change in the future under different scenarios (see for example, Dell, Jones, and Olken (2014) or Auffhammer (2018) for a discussion). Second, a common type of damage function is a fixed-effects model estimation to causally identify the impact of shocks on various outcomes (see for example, Dell, Jones, and Olken (2014)). Such estimations require ex-post panel data, which is not the use case of real-time monitoring.

Vulnerability functions estimate the effects of natural disasters on poverty and welfare by assigning probabilities of impact to individuals or households. The likelihood (and magnitude) of being affected by shocks is also called vulnerability (Doan et al. 2023). Households are vulnerable to climate-related and other types of shocks through a variety of characteristics. The ability to anticipate, cope with, adapt, and recover from the impact of a shock all determine the propensity of being affected (Hill and Porter 2017). Information on vulnerability, calculated using variables in baseline survey data, can be combined with information on exposure to a specific shock, usually through the household's location, to evaluate the risk and potential impact of specific shocks. The magnitude of impacts is typically determined by a calibration of parameters for different household characteristics, which rely on estimated effects from previous or similar shocks or on estimates from sectoral experts. The calibrated parameter allows to estimate a simulation model, which calculates changes in consumption from a baseline (pre-shock) survey—with information on household expenditure or consumption—to after the shock.

A rapid assessment of the welfare impacts of the 2022 floods in Pakistan exemplifies the use of damage and impact models. [After the devastating floods in Pakistan in 2022](#), the World Bank team worked quickly to estimate the potential impacts on poverty. To do so, they calibrated a parameter mirroring the intensity of the shock on household welfare. Intensity depended on a combination of how vulnerable the household was to the shock (determined using variables such as asset ownership, sector of employment, or household income) and how exposed the household's location was to the shock. For the case of Pakistan, this approach quantified that nine million people fell into poverty due to the. This approach allowed for a rapid assessment of welfare impacts within two weeks of the floods occurring.

Considerations Regarding Damage Functions

Vulnerability and damage functions rely on setting parameters, which can profoundly influence results. Probabilities of how much a household characteristic contributes to consumption declines carry uncertainty. If they are derived from estimates from other countries, an issue is external validity, that is, the extent to which estimates are valid for the country in question. Similarly, estimates from past shocks can also carry margins of error, if shocks were not exactly the same and effects are non-linear, or the way households are affected has changed due to other external factors (related to model stability). Expert opinions have the advantage that they are derived from individuals highly knowledgeable on the specific setting, but they are subjective and knowledge about a sector in one country might not pertain to another country. One option to strengthen the reliability of results is to provide a range of estimates using different parameters.

Models discussed here do not account for second-order effects. Damage functions can include CGE models, an extension that we do not incorporate here. Therefore, adjustments and adaptations that households make, and second-order effects, such as droughts leading to changes in prices or the productive capacity of households, are not accounted for. Including general equilibrium models is possible (see section 1.2), but requires more data (Roson and Satori 2016).

Vulnerability is usually defined as a state (0/1) rather than as an intensity. Vulnerability functions identify which households are more likely to be affected by a hazard. Based on characteristics such as sector of employment, type of assets, and consumption levels, a household is deemed to be vulnerable or not to shocks. This classification neglects heterogeneity between households, where combinations of characteristics could make households more or less vulnerable.

Resources

- [The Social Protection Stress Test Tool \(STT\)](#). This tool is commonly used to assess climate-shock impacts. It is based on modelling the impact of past shocks on consumption, and the probability distribution of hazards. The tool simulates how many people fall into poverty, given a range of shock magnitudes (World Bank 2021a).
- [The Unbreakable model](#). This model focusses on asset destruction and changes in income streams through natural disasters and covariate shocks (Hallegatte et al. 2016). It calculates the risk of each household to fall into poverty based on household characteristics. The model has shown that near-poor households are more likely to be pushed into poverty, for instance in the Philippines (Walsh and Hallegatte 2020). It also aggregates micro-level effects into macro-economic effects, such as aggregate economic and productivity losses, yielding a rich set of results for countries.



WORLD BANK GROUP