WPS7696

Policy Research Working Paper 7696

# Measuring Firm-Level Innovation Using Short Questionnaires

## Evidence from an Experiment

*Xavier Cirera*
*Silvia Muzi*

WORLD BANK GROUP
Trade and Competitiveness Global Practice Group
June 2016

Policy Research Working Paper 7696


## Abstract

Little is known about innovation in developing countries, partly because of the lack of comparable and reliable data. Collecting data on firm-level innovation is challenging because of the subjective definition of what determines an innovation, a problem that is exacerbated in developing countries where innovation is likely to be more incremental and less radical. This paper contributes to the literature by presenting the results of an experiment aiming to identify the survey instrument that better captures firm-level innovation in developing countries. The paper shows that a small set of questions included in a multi-topic, firm-level survey does not provide an accurate picture of firm-level innovation and tends to overestimate innovation rates. Issues related to framing explain some of the unreliability of innovation responses, while cognitive problems do not appear to play a significant role.


This paper is a product of the Trade and Competitiveness Global Practice Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at xcirera@worldbank.org and smuzi@worldbank.org.

# Measuring Firm-Level Innovation Using Short Questionnaires: Evidence from an Experiment[1]

Xavier Cirera Silvia Muzi

Keywords: Innovation, Measurement, Surveys, Firms, Development
JEL codes: 010, 030, 050

[1] Contacts: Xavier Cirera (Innovation and Entrepreneurship, Trade and Competitiveness, World Bank, xcirera@worldbank.org); Silvia Muzi: (Development Economics, Enterprise Analysis Unit, World Bank, smuzi@worldbank.org).
We would to thank Jorge Luis Rodriguez Meza and the participants in seminars at the World Bank and OECD for their suggestions and comments. We would also like to thank Xie Huang for excellent research assistance, Leonard Sabetti for his valuable and detailed work with the verification of the innovation descriptions, and Amanda Franziska Zarka for her comments. We thank the UK Department for International Development (DFID) for the financial support provided to DECEA for collecting data on innovation in developing countries.

# 1. Introduction

The critical role of innovation in economic growth and economic development is uncontested. Innovation is the engine of the Schumpeterian process of "creative destruction" where innovative firms and entrepreneurs continuously drive allocative efficiency and productivity growth (Schumpeter, 1942). A large empirical literature, mainly based on data on OECD countries, documents a robust positive relationship between firm-level innovation and productivity (Hall et al., 2009; Hall, 2011); innovation and employment (Harrison et al., 2008); R&D and productivity (Hall et al., 2010).

Yet very little is known about the extent and impact of firm level innovation in developing countries. The majority of existing studies are limited to either one country or a specific industry (Bogliacino et. al, 2012). Some exceptions have focused on the experience in Latin America, using cross-country analysis (Turriago, 2003; Hall e Maffioli, 2008), highlighting the differences with the experience of European countries (Raffo et al., 2008), or the linkages between innovation and productivity (Crespi and Zuniga, 2012; Zuniga and Crespi, 2013; Crespi et al., 2014). Outside Latin America the evidence is even more limited. Cirera et al. (2015) have analyzed the impact of innovation on productivity in a large sample of countries in Sub-Saharan Africa, South Asia, Eastern and Central Europe and the Middle East. Other studies include a recent analysis of the impact of corruption on innovation (Paunov, 2016) and a study on the use of the Internet as a driver for inclusive innovation (Paunov and Rollo, 2016). This lack of evidence has greatly undermined the ability to design appropriate and evidence-based policies to promote innovation and support economic growth in developing countries.

One of the main reasons for this lack of evidence is a lack of adequate data. Despite significant efforts in recent years by national and international agencies to implement innovation surveys in developing countries (Unesco, 2015; Nepad, 2014), innovation data are still scattered and suffer from several limitations. First, available innovation data are hardly comparable (Bogliacino et. al, 2012). Innovation surveys in developing countries are mainly conducted on a country-by-country basis without following a standardized methodology. Differences in survey instruments, sampling methodology and population of inference make comparison and benchmarking of innovation indicators across countries extremely challenging.

Second, although innovation in developing economies is likely to be different from innovation in developed countries - more incremental and less radical - most of the questionnaires used in innovation surveys are based on the Oslo or Bogota manuals (OECD/Eurostat, 2005 and Jaramillo et al. 2000) and seem to have some difficulties in capturing the relevance of innovations, to the extent that innovation

rates in developing countries appear to be larger than in OECD countries. This is likely to be explained by the subjective nature of some of the key innovation concepts. Respondents in different countries may have different understanding of what is considered as innovation, particularly when they are asked whether the firm has introduced a "new or significantly improved" product or process. In countries with higher prevalence of very incremental innovation, such as developing countries, it becomes even more subjective to judge how substantial or significant these innovation efforts are.

Third and more importantly, implementing innovation surveys is costly, especially in developing countries where the use of self-administered or web-based questionnaires is limited and data collection at the firm level heavily relies on face-to-face interviews. Comparability and reliability of the data collected are two key elements to consider when implementing innovation surveys. Measuring innovation is costly and difficult, and identifying the best instrument to collect innovation data is key from a cost-benefit perspective.

These challenges motivated the Enterprise Analysis Unit of the World Bank Group to conduct an experiment to test two different approaches to measure innovation in developing countries. The first approach, more economical, uses a short questionnaire embedded in a general firm-level survey, namely the Enterprise Surveys (ES). The second approach uses a standalone, longer questionnaire and it is significantly more expensive, although it provides more information.[2] The hypothesis to test is whether a cheaper approach could provide reliable higher frequency innovation data for key innovation inputs and outputs indicators by embedding a few questions in general firm-level surveys.

This paper summarizes the results of the experiment. Using a rich data set of 11,150 firms in 15 countries we quantify differences in innovation rates as measured by the short and long questionnaire and, more importantly, based on insights from the subjective measurement literature, we provide some empirical evidence on what determines these differences.

First we compare self-reported innovation rates in the short and long questionnaire across different types of innovation outputs (product, process, and marketing) and one measure of innovation input (R&D). Second, focusing on product and process innovation, we look at the description of the innovation introduced at the firm level to establish to which extent self-reported innovations are actual innovations. We then compare verified and self-reported innovation to determine whether the short or the long questionnaire is correct. Finally, as one of the main challenges of collecting data on firm-level innovation

[2] The standalone questionnaire was implemented as a follow up of the ES thereby reducing the need to collect data on other firm level variables that are helpful to understand the role and effect of innovation.

is the subjective nature of the innovation measurement, we examine the extent to which the different types of bias that occur when collecting subjective data play a role in innovation measurement.

Results show that in many cases self-reported innovations in the short and long questionnaire are different, even when the same person is interviewed in the two waves or when explanatory cards with examples or definitions to improve the comprehension of the innovation questions are shown. More importantly, for self-reported innovations that can be verified we find that the short questionnaire provides an incorrect answer in a significant number of cases. We find that issues of framing explain some of this unreliability, while cognitive problems do not appear to play a significant role. Overall, we found that some of the questions by which innovation is measured in current surveys are not appropriately designed to measure innovation in developing countries. This is especially important when using short surveys and, therefore, we recommend not to use short innovation modules with only a few questions in firm-level questionnaires that do not provide detailed verification questions.

The paper makes several contributions to the discussion on how to measure innovation. First, to the best of our knowledge this is the first paper that builds on the literature on subjective data to empirically analyze the potential sources of bias that may affect respondents' behavior in innovation surveys. Second, the paper makes use of a unique effort conducted by the World Bank Group to increase the coverage of innovation surveys. The data used are the most comprehensive set of comparable cross-country surveys on innovation in developing countries available so far. Finally, the paper reports on a unique experiment carried out to analyze methodological issues affecting data collection efforts and to assess the robustness of the Oslo manual framework in the context of developing countries.

The paper is structured as follows. The next section discusses the conceptual framework for the analysis. Section 3 presents the data used and the experiment conducted in data collection. Section 4 describes the extent of firm level innovation in the two surveys and analyzes the main differences. Section 5 empirically explores the determinants of these differences. The last section concludes.

# 2. Conceptual Framework

## 2.1 Measuring firm-level innovation outcomes

Innovation requires the transformation of knowledge capital or innovation inputs into innovation outputs, such as the introduction of new products or processes, increased quality of existing products or processes, marketing or organizational changes, and patented intellectual property. Firms invest in knowledge capital and innovation inputs in order to increase their capabilities to create innovative outcomes. In

addition to tangible innovation inputs- such as technology, equipment, and physical production facilities, innovation requires intangible assets such as human capital, scientific and creative capital and, more importantly, managerial and organizational capital. These inputs require in turn specific innovation activities depending on the degree of novelty and technology sophistication of the innovations introduced. Firms invest in training in order to increase their human capital available. In addition, firms can also invest in R&D, software and digitalization or copyrights, patents and licenses, and equipment in order to increase their scientific or innovative capital. Finally, innovation also requires organizational capital through adoption of new business models, design and prototyping or corporate alliances and networks. There is, however, uncertainty about the outcomes that can be achieved with the combination of inputs as innovation outcomes depend on each firm's ability, the sector and country context, and the enabling environment and policy framework.

When measuring innovation, the emphasis can be on measuring inputs and innovation activities and/or innovation outcomes. The earlier literature on innovation measurement focused on a specific set of innovation inputs that were more easily quantifiable, such as R&D or the technological intensity used. These early efforts were subsequently complemented with the framework provided by the Oslo manual, currently the main reference for innovation surveys, which mainly focuses on measuring innovation outcomes such as new products and/or processes, product and process improvements or patents at the firm level (OECD/EUROSTAT, 2015). A third generation of synthetic innovation indicators, such as the OECD STI scoreboard, was developed later on. These indicators combine innovation inputs and outputs in order to facilitate cross-country benchmarking and comparisons (OECD, 2015). However, to date, the Oslo manual framework, implemented originally through the Community Innovation Surveys (CIS) in the EU, has dominated innovation measurement in most countries, to the extent that even countries that do not conduct CIS have adopted the same questions at least in relation to the measurement of innovation outcomes.

One challenge in measuring innovation outcomes is the subjective nature of many of the questions used in the surveys. The Oslo manual defines innovation as "the implementation of a new or significantly improved product (good or service) or process, a new marketing method, a new organizational method in business practices, workplace organization or external relations". Most surveys use this definition to identify innovations, by directly asking firm managers and owners whether they have implemented any "new" products, processes, marketing methods, or organizational practices or "significant" improvements in existing ones in the last three years. In this context, the answer to this question is problematic since

innovation is self-reported and becomes a highly subjective concept, particularly when innovation refers to "significant" improvements. Therefore, the distinction between innovations and mere product, process, marketing or organizational changes becomes murky.

In recent years, several authors have advocated for a focus on knowledge capital assets rather than on innovation inputs or outcomes, leading to a renewed effort to capture investments in intangible assets using data from different sources (Corrado et al., 2005, 2006, and 2011; Hulten & Hao, 2012). In this approach, three broad categories of intangibles are identified which have some overlap with what is commonly considered as innovation-related activities (or capabilities): (1) computerized information (software and databases); (2) innovative property (R&D as well as intellectual property protection costs, architectural and engineering designs); and (3) economic competencies (brand names, firm-specific human capital and organizational capital). While this approach offers a better and broader measure of firms' capabilities, it poses a challenge as the information required cannot be obtained from existing innovation surveys. Moreover, it's important to highlight that measures of innovation assets are not equivalent to innovation outcomes, as inputs can be used more or less efficiently. Therefore, there is some uncertainty about the type and extent of innovation outcomes that can be produced by firms using their knowledge assets.

## 2.2 The problems with subjective measurement

As discussed above, one of the main challenges when measuring firm-level innovation outcomes is the subjective nature of the definition of innovation used in the survey questionnaires. While in the case of "new" products, processes, organization practices or marketing methods, the main challenge is to check the reliability of the self-reported outcome, an additional challenge arises when trying to capture the introduction of "significantly improved" products, processes, organization practices and marketing methods. The definition of "significantly improved" is highly subjective and can vary across individuals in the same country but also across countries.

Subjective data have strong limitations as measures of actual outcomes. In an influential study, Bertrand and Mullainathan (2001) show the likely bias when analyzing subjective data given their likely correlation with context variables. The authors conclude that while subjective variables can be used carefully as explanatory variables or to explain behavioral differences between individuals, models that use subjective data as dependent variables are likely to produce biased estimates.

The possible sources of bias when using subjective data are numerous, such as those that derive from cognitive problems as shown in Bertrand and Mullainathan (2001). A large amount of experimental

evidence shows that certain characteristics of the survey such as how the questions are structured in the survey, the ordering and the wording used, the format of the interview, or the quality of the translation, all significantly influence the outcome of survey responses. For example, whether the question is formulated in a positive way or in a negative way is more likely to translate into a positive or a negative answer. In addition, Schwarz (1999) describes how closed vs open response formats or the rating scale can also bias the response obtained.

Further sources of bias in subjective measures include context, cultural differences, and social desirability. The context in which questions are formulated is likely to influence answers, mainly due to the framing of the respondents. For example, Kaplan and Pathania (2010) show how perceptions-based indicators used to measure the quality of the business environment are highly correlated with the business cycle. Using data from the World Bank Enterprise Surveys (ES) the authors show how perceptions about the business environment worsen during periods of higher economic growth and, therefore, how changes in perceptions are not a good measure of changes in the business environment. Cultural differences are also relevant to how individuals understand and respond to survey questions. This is because aspects of cognition that affect response behaviors can differ across cultures (Johnson et al., 1997; Harkness et al. 2002). Village-level characteristics, such as ethnic heterogeneity and social participation, for example, have been shown to influence perception measures of corruption (Olken 2009). Furthermore, there is a large literature showing how social desirability affects the quality of subjective measures. One example is the reluctance of individuals to express subjective opinions perceived as non-desirable such as racial attitudes (Bertrand and Mullainathan, 2001).

A final important element that affects the bias in subjective measures is the reference or recall period (Schwarz, 1999). The time reference for the question or the recall period in the formulated question impacts the quality of the answer, and has been shown to impact both objective and subjective measures. Das et al. (2012) show that the length of the recall period, comparing a monthly recall with a weekly recall, had a large impact on reported morbidity, doctor visits, sickness, or reported use of self-medication in India. Bertrand and Mullainathan (2001) show that in some cases subjective answers can change over time due to the fact that the individual answering the question may not have established any preferred attitude.

## 2.3 Methodological problems in innovation surveys

The sources of bias identified in the literature on subjective measurement are particularly relevant for innovation surveys given the subjective nature of many of the questions used in innovation surveys. The

use of standard survey instruments in majority of innovation data collection in OECD countries has mitigated the likelihood of a different impact of cognitive issues like ordering or wording used on the quality of the data collected. However, methodological issues related to the respondent's understanding of innovation or to the type of survey instrument used have yet to be solved. The analysis of innovation and core economic indicators for EU countries, for instance, pointed out the existence of what is called the "Norwegian puzzle", which shows relatively lower innovation rates for Norwegian firms than what would be expected when comparing the Norwegian economy to that of other countries. The results of an experiment conducted by Statistics Norway suggest that this is partially explained by the survey instrument used. The experiment shows that innovation rates are significantly different when captured by a stand-alone innovation survey versus a combined R&D and innovation survey (Wilhelmsen, 2014). Similarly, an experiment conducted in Flanders region of Belgium shows that innovation rates are considerably and systematically higher when measured with a short questionnaire than when using a long questionnaire (Hoskens, 2015). While results are not clear as to which survey instrument is most accurate with respect to measured innovation rates, they clearly suggest that survey methodology can greatly impact the results.

In relation to context, cultural differences and social desirability there are several factors that can play a relevant role in how respondents understand and respond to innovation questions. Respondents are likely to have different views on what is considered as innovation depending on the context in which they live and operate. People in the USA, for example, are more likely to identify innovation with something new or unique and not to consider improvements as innovation. In Europe, on the contrary, people are more inclined to indicate both novelty and improvement as elements of innovation (Galindo-Rueda and Van Cruysen, 2015). Furthermore, self-reported innovation may be biased by social desirability that, in turn, might have a differential effect on respondents from different cultural backgrounds (Johnson & van de Vijver, 2003). As innovation is commonly associated with increased productivity and growth, it may perceived as a desirable outcome. Therefore, respondents to innovation surveys may be inclined to over-report innovations when follow-up questions aiming at ascertain the accuracy of the answer are not asked and, furthermore, the tendency to over-report innovation may vary across countries.

Finally, methodological issues can also have an impact on the measurement of innovation across firms of different size, as documented in a recent study conducted in Poland. The study, which compares innovation rates of micro and non-micro firms, shows higher innovation rates for micro firms as compared

to small and medium firms (Rozkru, 2015). This counterintuitive result may be explained by the fact that micro firms may be more likely than larger firms to report improvements as innovation.

In the following sections we attempt to contribute to the literature on the impact of survey methodology on innovation measurement by shedding some light on the importance of problems typical of subjective data when measuring innovation in developing countries.

## 3. Data

This paper draws on Innovation Enterprise Surveys data collected between 2012 and 2014 on a representative sample of the formal (registered) non-agricultural, non-mining private sector in Sub-Saharan Africa and South Asia. Fifteen countries are covered by the surveys: 11 in Sub-Saharan Africa (DRC, Ghana, Kenya, Namibia, Nigeria, Sudan, South Sudan, Tanzania, Uganda, Malawi and Zambia) and four in South Asia (Bangladesh, India, Pakistan, and Nepal) (see Table A2 in the Appendix for more details). The data were collected by the Enterprise Analysis Unit of the World Bank Group as a part of a DFID-funded project, which aimed to increase the coverage of innovation surveys in developing countries while assessing at the same time the robustness of the Oslo Manual framework.

The project followed a two stage approach to data collection. In the first stage, a short innovation section (short questionnaire) consisting of 9 questions based on the CIS-type of survey was incorporated into the Enterprise Surveys (ES). The ES is a multi-topic firm-level survey that collects data on firms' characteristics, balance-sheet, and firms' experience of the business environment.[3] In the second stage, an in-depth innovation module (long questionnaire) was administered to randomly selected ES respondents. The questionnaire used in the innovation module is also based on the Oslo manual but expanded to capture more details on the innovation introduced and to integrate other important questions on organization and management practices, use of ICT, and innovation inputs.[4] The follow up survey was submitted to firms independently from their answers to the innovation section in the ES. Whenever possible the same respondent was interviewed in the two rounds of data collection.[5]

[3] More information on the Enterprise Surveys are available on the website http://www.enterprisesurveys.org/

[4] Methodological tests and pilot field work were conducted in preparation to the in-depth innovation module. The methodological tests was conducted in Ethiopia, Rwanda and Zimbabwe and included preliminary innovation surveys complemented by focus groups and semi-qualitative interviews. The in-depth questionnaire also benefited from inputs from experts on innovation in the National Science Foundation (NSF), WB and IDB.

[5] The World Bank Group Enterprise Analysis Unit also collected data on innovation in Europe and Central Asia (ECA) and Middle East and North Africa (MENA) for a total number of 49 countries with innovation data. Innovation Enterprise Surveys in ECA and MENA are also collected following a two stages approach. The methodology for the

The main advantage of collecting innovation data as part of the ES is that data are fully comparable across countries, allowing benchmarking and comparison of indicators. Moreover, innovation indicators can be related to a broader set of firm-level indicators covering firm's performance, job creation and experience of the business environment. The ES are conducted by using a global methodology, which includes the same universe of inference, same sampling methodology – stratified random sampling –along with a common questionnaire. The sample for each country is stratified by firm-size, sector of activity and location within the country. Weights are provided in the survey to ensure that the sample is representative of the formal (registered) non-agricultural, non-mining private sector of the economy.[6]

The objective of the project was to test the reliability of the answers provided in the short questionnaire in order to determine if a short set of questions on innovation can be incorporated into standard firm-level surveys. As suggested above a challenge in measuring innovation, using CIS-type of surveys is the great subjectivity of what constitutes an innovation. In the case of the short questionnaire this problem is exacerbated by a lack of follow up control questions that can provide some details to guide the veracity and accuracy of the answers. Additional methodological issues, like respondents' reactions to combined questionnaires vis-à-vis to single topic questionnaires, should be considered when choosing the most appropriate instrument for innovation surveys. The project also incorporated an additional test to determine the extent to which cognitive problems play a role in biasing respondents' answers. Half of the respondents in the short questionnaire were shown explanatory cards with examples or definitions to improve the comprehension of the innovation questions.[7]

Four types of innovation outcomes - product, process, marketing and organization - and one innovation input – R&D - were considered in both surveys. Following the Oslo manual, product innovation is defined as the introduction of new or significantly improved product (goods and/or services) and process innovation is disentangled into three components: methods of manufacturing goods or offering service ("methods" from now onwards); logistics, delivery, or distribution methods for inputs, products, or services ("logistics" from now onwards); supporting activities such as maintenance systems or operations for purchasing, accounting, or computing ("auxiliary" from now onwards). For product and process

follow-up questions, however, is slightly different as innovations questions are administered only to firms identified as innovators based on the ES.

[6] We note that Enterprise Surveys cover manufacturing as well as services sectors but certain services such as education and health are not covered. Moreover, the primary sector, which encompasses agriculture, mining, forestry, etc., is also excluded from the survey.

[7] The experiment of using show cards on half of the sample was conducted in Ghana, Kenya, Tanzania, Uganda, and Zambia.

innovation, the two survey instruments are fully comparable as the same questions are asked in both surveys. The long questionnaire, however, includes some additional follow-up questions designed with the purpose of collecting additional evidence on reported innovation while systematically probing the answers to the standard yes/no questions. Follow-up questions include, a question on the year in which the innovation was introduced as well as open questions where respondents are asked to explain the innovation introduced and how this is different from previous products or processes in the firm. The description variables are critical in our analysis since they allows us to verify the innovation implemented. As a result, for product and process innovation our dataset contains three measurements: response in short questionnaire, response in the long questionnaire and response in the long questionnaire verified based on the description provided.

For organization and marketing innovation, the short questionnaire includes the standard questions based on the Oslo manual,[8] while the in-depth innovation survey disentangles different aspects of organizational and management innovation and different marketing strategies in which firms may innovate. Answers to these more detailed questions are mapped to the more general yes/no question included in the short questionnaire. Finally, four questions about R&D are asked in the long questionnaire, covering firms' engagement in internal and external R&D and the associated costs. Those questions are mapped to the yes/no question on R&D included in the short questionnaire. The reference period is the same in both survey instruments.

The following analysis focuses mainly on product and process innovations as these can be verified using the additional information in the long questionnaire to corroborate the reliability of the answers provided. We also provide some information on marketing innovation and R&D.[9] Data cover about 22,000 firms operating in the manufacturing and services sectors in the short section in the ES and 11,150 firms in the innovation follow-up survey (see Table A2 in the Appendix for more details on the sample by country).

[8] "During the last three years, has this establishment introduced any new or significantly improved organizational structures or management practices?" for management and "During the last three years, has this establishment introduced new or significantly improved marketing methods?" for marketing.

[9] Organizational innovation is not included in the analysis that follows. The section of questionnaire dealing with organizational innovation is very different compared to the other section and the analysis of the results would require a separate approach.

# 4. Measured innovation outcomes in the short and long questionnaire

## 4.1 Innovation rates in the two surveys

Results from the two stages of data collection were expected to show some variation given the differences in the two survey instruments and the subjectivity of the questions asked. The extent of these differences, however, is surprisingly high. Overall, in the full sample more firms reported innovation in the short questionnaire as compared to the long questionnaire in all types of innovation but marketing (Figure 1).[10] The magnitude of these differences varies considerably, ranging from 1 percentage point for R&D to 14 percentage points for process innovation. Moreover, high differences in reported innovation rates are found in all three components of process innovation with the highest difference being in auxiliary activities and logistics (14 percentage points difference) and the smallest in methods (9 percentage points).

Despite the lower innovation rates in the long questionnaire, the proportion of firms that reported innovations are still higher than expected: 19 percent of firms in the full sample engaged in R&D activities, 39 percent of firms reported product innovations and 45 percent of firms reported innovations in processes. Marketing innovations showed an opposite pattern with fewer firms reporting innovation in the short questionnaire (47 percent) compared to the long questionnaire (68 percent). The fact that in the long questionnaire marketing methods were disentangled into several components, each of them with a specific yes/no question, can explain this opposite pattern. In this case the problems with subjective data measurement apply for each of the components in marketing methods and get, therefore, magnified in the complied measure.

The existence of considerable variation in self-reported innovation is also confirmed at the country level. As Figure 2 shows, differences in reported innovation are sizable across types of innovations and countries.[11] However, the magnitude and direction of these differences vary widely and no country shows consistently lower or higher differences across types of innovation. Despite this heterogeneity, some patterns can be identified. First, with the only exception of Bangladesh, India, South Sudan, and Zambia, innovation rates in product and process are higher when measured by the short questionnaire. Moreover, differences in innovation rates for these two types of innovation are high in most of the countries. Second,

[10] Innovations rates are computed by using sampling weights. Overall averages are computed as simple average of weighted country-level averages.

[11] A simple test of the equality of the means shows that we can only accept equality of the mean between the short and long survey measure for product, process and marketing innovation in DRC, marketing innovation in Kenya and R&D for Ghana, Nepal and South Sudan.

R&D is the type of innovation characterized by the smaller variation between the two surveys. Although this may be in part driven by the smaller number of firms reporting R&D, the lower differences can also be an indication of a better familiarity of respondents with the concept of R&D investments and therefore smaller measurement errors due to cognitive or framing problems. Finally, innovation rates for marketing methods are higher in the long questionnaire for almost all countries in the sample for the same reason discussed above.

Despite the fact that innovation rates are lower in the long than in the short questionnaire, they are still very high, particularly when the distinction between innovation at the technological frontier and adoption (and adaptation) of existing technologies is not considered. Almost half of the countries in the sample exhibits product and process innovation rates of 40 percent or higher. Reported innovation in marketing is even higher with the innovation rate lower than 50 percent only in two countries. Lower, but still high rates are reported for R&D, ranging from 7 percent in Nepal and Pakistan to 26 percent and 28 percent in Namibia and Malawi and to an even higher 58 percent in India (innovation rates by country and type of innovation are presented in table A3).

Given the existence of considerable differences in innovation rates across countries and the high innovation rates reported in both surveys it is important to verify whether measured innovation rates are accurate. As described above, the long questionnaire provides a description of the new or significantly improved product and processes introduced for all firms. This allows us to verify innovation information and to build a "clean" innovation variable.[12] The cleaning and verification process was possible for around 48% of firms for product innovations and 53% of firms for process innovations (a description of how the cleaning exercise was implemented is provided in the Appendix). The remaining firms reported no innovations in their products and processes. The implication of this, as we will explore below, is that the correctness of the reply for non-innovators cannot be verified. This is a problem in the analysis of the determinants of discrepancies in innovation rates since some of the elements that may bias the responses

[12] The following analysis is based on the sample of 11,150 firms for which this information is available. Follow-up questions to be used for the verification and cleaning process were not asked for marketing innovation or R&D. As such the innovation rates based on these two measures cannot be verified. An additional verification was conducted by using the year in which the innovation was introduced to make sure that all reported innovations were introduced in the correct reference period. The question asked whether new or significantly improved process or product were introduced during the past three years but in a few cases respondents reported innovation introduced before that period.

for innovators also apply to firms that report themselves as non-innovators. For example cognitive problems, framing or survey fatigue may bias some responses by non-innovators.

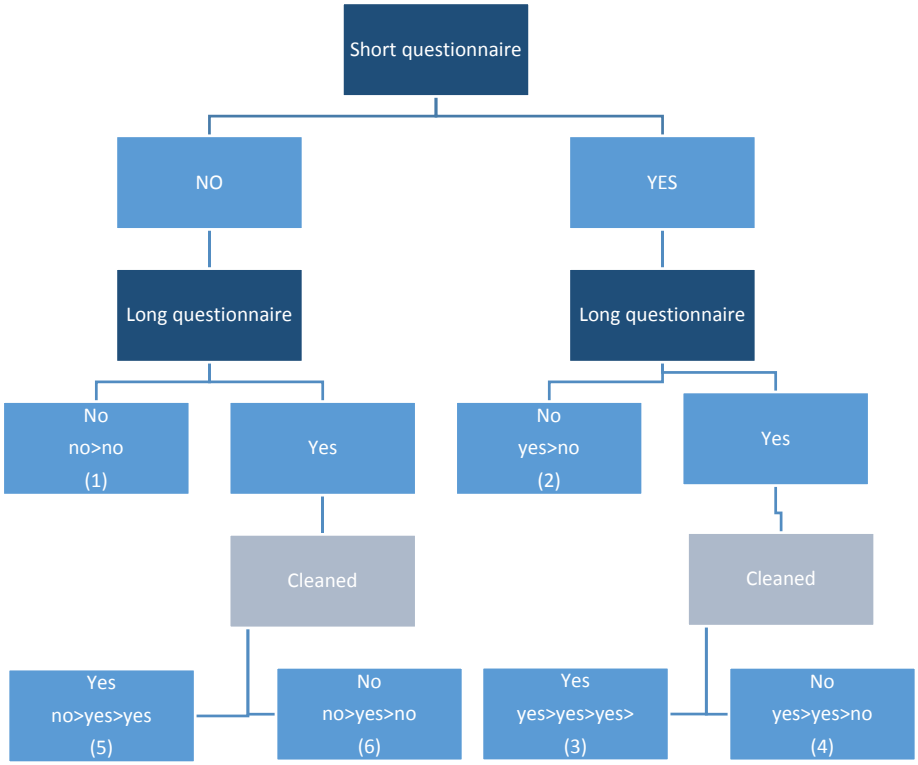
Figure 3 compares combined product and process innovation rates as measured in the short questionnaire, the long questionnaire and the cleaned long questionnaire. In all countries the cleaning exercise reduces innovation rates compared to the long questionnaire, although by different magnitudes. The smallest differences are in Namibia (4 percentage points), Nepal (4 percentage points), and Malawi (6 percentage points) while the highest are in Bangladesh and Sudan, where innovation rates dropped by 20 percentage points and 33 percentage points respectively. Innovation rates measured by the long clean questionnaire are also smaller than in the short questionnaire for all countries with the only exception of Bangladesh. When measured using the clean long questionnaire, combined product and process innovation rates range from 6 percent in Sudan, 15 percent in Pakistan, and 21 percent in Nepal to 78 percent in Bangladesh and 84 percent in India.

## 4. 2. Which survey provides a more accurate measurement of innovation?

Which questionnaire better captures innovation? Is a long questionnaire a good -or better- instrument to measure innovation in developing countries? The assumption is that a long questionnaire is more likely to capture an accurate picture of innovation as it gives greater information on the context and asks control questions that can potentially increase the accuracy of reported innovation. However, despite these reasonable assumptions, the subjective and self-reported nature of innovation information makes it possible for firms to overestimate or even to underestimate innovation efforts also during in-depth, single topic questionnaires. Moreover, long stands-alone surveys can contribute to increase survey fatigue opening up space to underreport innovation in order to reduce the burden of the interview.

Insights on these possible scenarios can be obtained by looking at respondents' answers in the experiment. Figure A decomposes the different combinations of responses to innovation in the surveys, including the short, long and long-clean questionnaire. Six categories of firms are identified: firms that replied no to both short and long questionnaire (no>no); firms that replied yes to the short questionnaire and no to the long questionnaire (yes>no); firms that replied yes to both questionnaires and for which the answer was confirmed in the long-clean questionnaire (yes>yes>yes); firms that replied yes to both questionnaires but for which the answer was not confirmed (yes>yes>no); firms that replied no to the short questionnaire and yes to the long questionnaire, with the yes confirmed after verification (no>yes>yes); and firms that replied no to the short questionnaire and yes to the long questionnaire, with the yes not confirmed in the long-clean questionnaire (no>yes>no).

Figure A – Reported innovation in the short, long and long-clean questionnaire




The percentage of firms in each of the six categories by country is presented in Table 1. Column (7) shows the percentage of answers that could be verified, which is equal to the share of reported innovators in the long questionnaire, since these are the firms for which the description of innovations is available. In total 70 percent of the answers were verified with the share of verified answers ranging from less than 40 percent of firms in Nepal, Pakistan, and Sudan to almost all firms in Bangladesh and India. Column (8) shows the percentage of firms that provided a correct and consistent answer in the short questionnaire. On average, 66 percent of the responses were correct, which means that in 34 percent of the cases innovation was not measured correctly when using the short questionnaire. In some countries such as India, South Sudan, and Zambia the short questionnaire provided a correct measure of product and process innovations for more than 75 percent of firms, but in countries like Nepal or Sudan the short questionnaire was successful in measuring innovation in less than 50 percent of the sample. Note, however, that this is based on the assumption that the firms answering no in both surveys (1) or reporting

no innovation in the second questionnaire (2), are not providing biased answers; these cannot be verified, and therefore the share of correct answers is an upward estimate.

As seen above, in some countries there are considerable differences between the short and the long questionnaire (column 9). Which measure is correct in these cases? Column (5) shows the cases where the change of response in the long questionnaire could be verified as correct while column (6) reports the cases in which the long questionnaire introduced measurement errors. On average, 8% of cases that changed their answer were able to report innovation correctly (column 5): that is, the long questionnaire gave them the opportunity to correct an initial non correct answer. On the other hand, as column (6) suggests around 2% of cases that did the same misreported innovation: that is the long questionnaire open to door to a mistaken answer. Finally, the last column (10) shows the percentage of cases e that were re-classified based on the description provided in the long questionnaire, around 13% on average. These are clear cases where there are cognitive problems in understanding what constitutes an innovation.

In order to analyze more in depth the reasons for misreporting innovation we looked at the basic characteristics of firms in the six groups identified based on the different combinations of responses to the surveys. In Table 2 for each group we tabulate basic firm's characteristics and in Table 3 we present indicators of the extent to which each group engages in knowledge activities, an important predictor of innovation outcomes.

It is possible that more established and sophisticated firms, like large, old, exporting firms or firms with foreign ownership, may be more familiar with the concept of innovation and, therefore, less likely to misreport innovation. The same may be expected for manufacturing firms, as it can be easier for these firms, than for firms operating in services, to identify a product innovation and to disentangle product from process innovation. On the other hand, it's possible to argue that small and young firms are the more dynamic and innovative firms in the playing field. They can be expected to be more familiar with the concept of innovation and, therefore, less likely to report discrepancies between the two surveys.

The first four rows of Table 2 and Table 3 represent firms that are categorized as non-innovators; the first two are classified as non-innovators in the long innovation questionnaire and the third and fourth are re-classified as non-innovators after verifying the information about the self-reported innovation. First, there are remarkable differences between the first group (no>no) and the other groups in terms of firms' characteristics and innovative efforts. Non-innovators that provided a negative response in both surveys are smaller, younger, less likely to be exporters and invest less in knowledge inputs than all the other

categories. This suggests that this no>no group is likely to be a good representation of non-innovators. Firms that answered yes in the first survey and no in the second (yes>no) present more mixed characteristics. They are bigger and more likely to export compared to the no>no group. However, as firms in the no>no group, they are less likely to engage in knowledge activities compared to the other groups.

While groups in rows 3 and 4 (yes>yes>no and no>yes>no) are different from each other in terms of firm size and likelihood of being exporters or to have foreign participation in ownership, they do have a lower incidence of knowledge activities compared to the group of firms that are innovators –rows 5 and 5-, with the exception of the group that was reclassified as a non-innovator after having a positive response in both surveys.

So do non-innovators in the survey look like non-innovators? In general and looking at knowledge inputs as a predictor of innovator, it is likely that most firms labelled as non-innovators are in fact non-innovators; although there is a small grey area of about 10 percent of firms – the y>y>n group -where some firms with similar characteristics and incidence on participating in knowledge investments can be innovators or non-innovators. Firms' characteristics, however, seem to be a less straightforward predictor of innovator, apart from the no>no group.

Finally, even assuming that most firms that are classified as non-innovators are in fact non-innovators, the capacity to predict innovation using the short survey is limited, especially for services. Around 31 percent of firms in manufacturing (rows 2, 3 and 6), 41 percent in wholesale and retail and 36 percent in other services are misclassified in the short questionnaire.

Interestingly, the accuracy of responses did not increase even when a show card with detailed examples of innovations (i.e. what was intended as product, process, or marketing innovation) was provided to respondents during the first wave of data collection (Table 4). Contrary to expectations, the distribution of firms by different groups is the same for firms that were shown and not shown explanatory show cards. The rate of correct short questionnaires with no-show cards (rows 1, 4 and 5) is 63.22% and with show card is 62.47%.

Overall, the results show very large discrepancies when trying to measure innovation in both surveys, and important mismeasurement when using the short questionnaire. These findings are based on purely descriptive statistics. The next section tests the robustness of these results and dig more in-depth into the determinants of the identified discrepancies.

# 5. The empirical analysis

## 5.1 Explaining differences in responses between the two surveys

Based on the literature on subjective measurement and on the role of methodological issues on the measurement of innovation, we identified three sets of elements that may affect respondents' answers to innovation questions (Table A1):

- Cognitive problems due to different understanding of innovation and quality of the interview
- Respondent framing - both firm characteristics and context in which firms operate
- Recall period – time elapsed between the two interviews conducted for the experiment

Regarding cognitive problems, there are several factors that can play a critical role in the way respondents answer innovation questions. First, familiarity with the context of innovation is likely to be associated with accurate responses. This is represented by exposure to innovation activities as measured by whether or not a firm is currently engaged in or has abandoned innovation efforts in the past. Having additional information on what is considered as innovation is also likely to be associated with more accurate responses. This is captured by the use of explanatory cards with examples of innovations during the interview. The level of education of the manager or the years of experience of the respondents working with the firm are also likely to influence the response and make it more accurate. Finally, the gender of the respondent may play a role, as female and male respondents may be characterized by a different response behavior. The second set of issues that can affect the responses is linked to the quality of the interview. Quality of the interview is proxied by: whether the person interviewed in both surveys is the same; the number of interviews per supervisor -the higher the number of interviews per person the more difficult to control quality; and the accuracy of the answers provided, measured by the perceptions of the interviewer about the quality of the interview.

Regarding context and framing, the accuracy of answers is affected by survey fatigue that can be measured by several proxies. A first proxy is represented by the burden of the business environment, as measured by the number of meetings with tax officials. Having frequent meetings with tax officials can be a burden for firms as it reduce the time dedicated to productive activities. Therefore, as the number of meetings with tax official increases, firms may be more inclined to reply to surveys with less accuracy in order to quickly go back to productive work. Moreover, the length of the interview during the first wave of data collection is a good proxy for survey fatigue. Having already replied to a long interview could be an incentive to reply no to innovation questions in a follow-up interview and reduce the time devoted to

the survey. Also related to context and framing, respondents and managers have different attitudes to reporting "true" information; we proxy this attitudes with a dummy measuring whether the firm reports having been exposed to the payment of bribes. The direction of the effect is ambiguous. On the one hand a manager that is more likely to report having been exposed to bribes can be thought of as more likely not to provide true information, because more inclined to circumvent the system. On the other hand, one can also claim that reporting having been exposed to the payment of bribes, as opposed to paying bribes but not reporting it, is an indication of being inclined to provide accurate answers during the interview. Third, firms' characteristics may impact response behavior. Larger, older, foreign owned, and exporting firms may be more likely to provide accurate information due to reputational issues. Also, it is interesting to analyze the performance context of the firm since firms that are expanding (contracting) could have less incentives to provide inaccurate (accurate) information since they can show potential investors and clients a good performance. This is represented by the annualized growth in employment over a three years period. Finally, to explain the differences in answers between the two surveys it is important to consider the recall period, measured by the number of days elapsed between the implementation of the two surveys. The hypothesis is that the longer the time between the surveys, the lower the anchoring effect, as it is more likely that respondents forget the response provided in the first survey.

To empirically verify whether these elements play a role in explaining respondents' behavior we regress the differences in the answers between the two surveys on the variables identified above. We first use a probit model in which the dependent variable takes the value of 1 if the answer in the two surveys is different and 0 otherwise. However, the incentives that play a role in providing different answers in the two surveys maybe be different in yes to no answers –that is, yes to the short questionnaire and no in the long questionnaire, as compared to no to yes answers. In order to capture these different incentives we also estimate a multinomial logit model, where we model three different outcomes. Specifically, we consider the outcomes for $Y_i$ equal to 0, 1 and 2; where 0 represents no difference between the two surveys, 1 represents that firm i reports innovation in the short questionnaire and no innovation in the long questionnaire; and 2 when firm i reports no innovation in short questionnaire and innovation in the long questionnaire. Therefore $Y_i > 0$ represents discrepancy between questionnaires.

Table 5 shows the results of estimating the probit model with country and 2 ISIC digit industry fixed effects. Columns (1) and (2) show the results for the full sample, while column (3) and (4) are the estimates for the group of countries in which show cards where used. The time interval that passes between the two surveys and the number of interviews per supervisor increases the probability of having different

answers in the two surveys; as expected, interviewing the same person reduces the probability for discrepancy. The firms' size and ability to be a potential innovator are also negatively correlated with discrepancies in the answers provided in the two surveys. Interestingly, reporting the payment of bribes was shown to be negatively and significantly correlated with discrepancies, supporting the hypothesis that more candid respondents that have no reserve in reporting the payment of bribes are also more frank in answering the innovation questions.

Firm size is negatively associated with the probability for discrepancy probably because larger firms are more likely to know about innovation activities. This possibility seems to be confirmed by the fact that this significance disappears in the sub-set of countries where show cards were used to show examples of innovation. In the sub-set of countries in which show cards where used only the significance for being a potential innovator, time in between interviews, and reporting the payment of bribes holds. Finally, the use of show cards does not have any role in explaining discrepancies.

Results of the multinomial logit estimation are reported in Table 6. The coefficients need to be interpreted in relation to the baseline category of no discrepancy between the surveys. The columns labeled #a show the coefficients for the cases where the short questionnaire was answered with a yes for product and/or process innovation, while the long questionnaire provided a negative response. The columns labeled #b represent negative answers in the short questionnaire and positive in the long questionnaire. Columns 3 (a and b) and 4 (a and b) estimate the model only for the countries that used explanatory cards in a subset of the interviews.

Regarding cognitive issues or understanding of innovation concepts, firms that carry out at some point an innovation, even if it was abandoned, are less likely to change their mind from yes to no, and more likely to change their mind from no to yes. This is likely the result of a better understanding of the concept of innovation, but it might be also due to the fact that there is high correlation between potential and actual innovation. Neither the manager years of experience nor the years of experience of the respondent in the short questionnaire matters for explaining discrepancies; with the exception in the last case of changing their mind from no to yes. Further, while it was reasonable to find that the manager's experience was not a significant indicator, given that the manager was not always interviewed for each firm, the finding about the respondent's experience with the firm is somehow unexpected. The assumption is that a longer experience of the respondent with the firm would imply lower likelihood of different responses in the two surveys, at least if the difference is considered as the result of lack of knowledge about firm's innovation activities, although not always the same person is interviewed in both surveys. When the same person is

interviewed in both surveys, this reduces the probability of a discrepancy but only for the yes-no combination. More surprisingly, the introduction of explanatory cards does not affect the probability of observing a discrepancy. This suggests that overall most of the discrepancies are not likely to be the result of lack of understanding of innovation concepts, and other factors may be at play.

One of these factors is the recall period. The longer it takes for the firm to be re-interviewed, the more likely are discrepancies to arise. Another important factor to increase the likelihood of discrepancy, at least for the combination yes-no, is the number of interviews for which the same individual supervises over as a proxy for the quality of the enumerator interview. Interestingly, the perception of the enumerator about the quality of the first interview is not a good predictor of observing discrepancies.

Regarding the context of the firm and potential survey fatigue, those firms that have payed bribes are less likely to have a no-yes discrepancy across surveys. As we saw in table 1 in majority of cases the change from no to yes was confirmed as innovation. The "no" answer provided in the short questionnaire may have been then originated by the desire to protect the innovation or by a reluctance to share correct information if considered as sensitive. In this case one potential explanation is that firms that are candid about paying bribes are more open also to disclose innovation and therefore less inclined to reply no in the short questionnaire. On the other hand, firms experiencing positive growth in their performance represented by employment growth, appear more likely to a no-yes discrepancy. The average number of meetings with officials as a proxy for survey fatigue is not statistically significant explaining these discrepancies. Finally, large firms and foreign firms, (in one specification) are more likely to have one type of discrepancy (no>yes).

## 5.2 Explaining the quality of the responses in the short questionnaire

One important element to consider when interpreting these results is that in a significant number of cases that answered yes in the second questionnaire are then reclassified after the verification process as non-innovators. Therefore, looking at the discrepancy between surveys is not enough to assess the accuracy of the answers provided in the short questionnaire. For this, we need to look at the differences between the answers in the short questionnaire and the cleaned responses in the long questionnaire.

Table 7 and 8 show the estimates of the probit and multinomial logit model for the determinants of providing an inaccurate response in the short questionnaire. Since we are focusing on the ability of the short questionnaire to measure accurately innovation we focus on the potential determinants of errors in the first interview. As such, the time elapsed between interviews or whether the same person

responded in both interviews is no longer relevant to explain the errors. In Table 8, columns 1a, 2a, 3a, and 4a represent cases where there was conflicting information regarding the innovation status of a given firm. Specifically, there was an error in answering "yes" for innovation in the short questionnaire, given that the firm resulted as non-innovator when the answer was verified using the description of the reported innovation. Finally, columns 1b, 2b, 3b, and 4b represent the probability of the reverse error.

Regarding cognitive issues, the results suggest that these factors do not explain the likelihood of erroneously reporting innovation. Neither the experience of the manager or the respondent, nor the use of explanatory cards, have an impact on the probability of erring. Only familiarity with the concept of innovation, here proxied as the firm being a potential innovator, affects the probability of erring in the short questionnaire. However, the coefficient is positive and statistically significant, which is contrary to what we would expect. This suggests that factors other than lack of understanding are likely to explain these errors. Interestingly, the probability of providing an inaccurate response in the short questionnaire is negatively associated with the respondent being female. In particular, female respondents are less likely to answer yes in the short questionnaire when there is no innovation after the verification process.

In terms of the quality of the interview, the estimates suggest that for some specifications there could be an issue with the enumerator; the more interviewers are supervised by the same supervisor the higher the likelihood to collect an inaccurate response. Again, the quality perceived by the enumerator is not a statistically significant predictor of inaccuracy in the response.

Table 8 also shows that the impact of some of the framing variables explaining the inaccuracy of responses is also limited. Survey fatigue represented by the number of days spent by the firm dealing with tax issues did not show to be statistically significant. Interestingly, as discussed above, whether the firm reports the payment of informal gifts as a measure of the trustworthiness of responses reduces the probability of an inaccurate response in the no-yes. Finally, expanding firms are more likely to answer negatively in the short questionnaire when they are *de facto* innovators. This may indicate some desire not to be identified as innovators.

Firm characteristics are also not a good predictor of inaccurate responses with the exception of size and foreign ownership. Larger firms and foreign firms are less likely to provide a negative response when they are innovators. Larger firms in the sub-sample of countries where show cards were used are also less likely to report innovation that is not confirmed after the verification process, reinforcing the conclusion that the short questionnaire provides a less accurate measure of innovation for smaller firms. Age or exporter status do no play any role in explaining inaccurate responses.

One potential caveat of the estimates above is that only those firms that provided a positive response in the long questionnaire could be verified, since verification was conducted on the basis of the description of the innovation implemented. Although as we saw in the previous section, the characteristics of the firms that provided a negative response in both surveys are those of a likely non-innovator, especially given the lack of investments in knowledge capital; some of the firms in the yes-no outcome could be in fact innovators. One solution to this problem would be to estimate the econometric model only for the sample of firms for which innovation can be verified. However, one potential problem in doing this is sample selection bias, since the sample would then be non-random and it is reduced to those firms answering yes to innovation in the second questionnaire. If an unobserved variable explains the probability of answering yes in the second questionnaire but it is also correlated with the probability of providing an inaccurate response, then our estimates are biased.

In order to correct for the sample selection bias and estimate the model for the sample containing cases where innovation can be verified we employ a Heckman (1976) correction model and use a Heckit estimator. The intuition is that we can correct for the sample selection by using covariates that are strongly correlated with the selection in the sample but not with the outcome. In our case, in addition to firm characteristics, we use whether the firm invests in different types of knowledge inputs to help predict selection, since these are innovation inputs that should make more likely for the firm to provide a positive response, while likely to be uncorrelated with the decision to provide an inaccurate response.

Table 9 shows the results of estimating the Heckit model. Even columns estimate the probability of providing a positive response in the long questionnaire, and odd columns estimate the probability of providing an inaccurate answer in the short questionnaire - using also the inverse mills ratio, the ratio of the probability density function to the cumulative distribution function in the first stage, to control for potential sample selection bias and the potential effect of the unobservable covariates.

The knowledge input variables used in the first stage are statistically significant predictors of answering affirmatively in the long questionnaire, with the exception of purchase of knowledge such as patents or trademarks, which tend to be less common for innovation activity in developing countries. The inverse mills ratio coefficient, lambda, is positive and statistically significant suggesting a potential upward bias of the estimates if there is no control for sample selection bias.

Regarding the coefficients of interest, cognitive issues do not seem to play a role in explaining inaccuracy in the short questionnaire response. Again, the results confirm the lack of impact of explanatory cards. Also, the experience of manager or respondent does not seem to matter. Once we control for the

potential selection bias, having familiarity with the concept of innovation and the gender of the respondent are not significant predictor of providing a wrong answer in the short questionnaire. The main results suggest that a lack of quality in the interview, represented by the number of enumerators supervised by the same supervisor, increases the likelihood of inaccuracy; whether if the firm pays informal payments decrease the likelihood of inaccuracy. Furthermore, size is inversely correlated with inaccuracy with larger firms being more likely to correctly report innovation in the short questionnaire. Other unexplained factors appear to play an important role in explaining such inaccuracy.

## 6. Conclusions

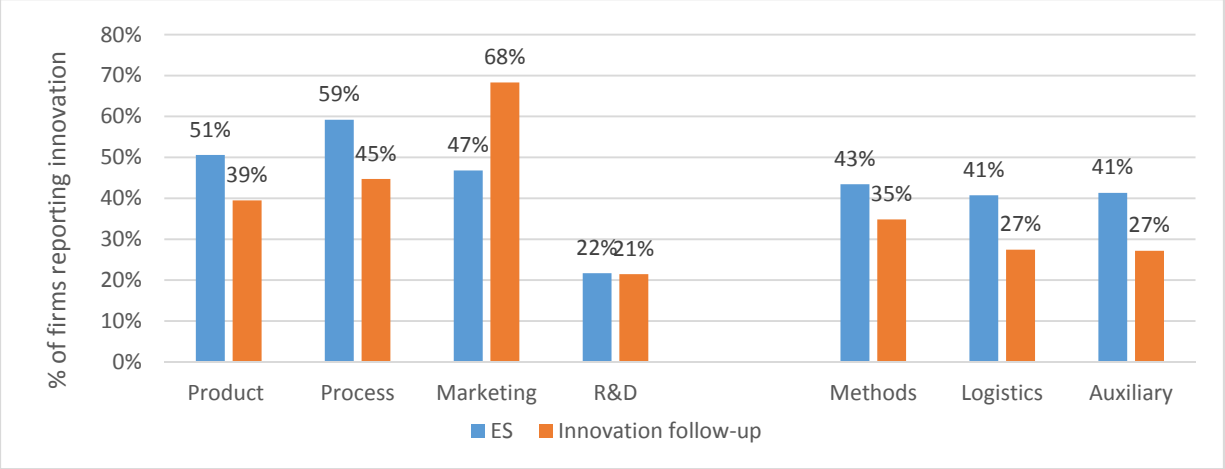
This paper shows the results of a survey experiment aiming to identify the best survey instrument for measuring innovation in developing countries. The results from the experiment are clear: a few short questions in a more general firm-level survey do not provide an accurate picture of firm-level innovation activity. This result confirms the findings of studies conducted for more advanced countries such as Norway or Belgium that suggest that shorter surveys tend to overestimate firm level innovation rates. More context to the questions is likely to be needed; although the innovation rates found in the longer questionnaire, which attempts to provide more context, also requires substantial cleaning and are unreasonably high given the level of development of the countries in the study.

The paper also tries to understand empirically what factors could explain the inaccurate responses to the short questionnaire and the fact that a significant number of firms provide different responses to the same question across questionnaires; sometimes even when the same respondent answered the questionnaire. Results show that lack of knowledge does not seem to play a key role in explaining inaccuracies, as neither the use of explanatory cards nor the experience of the respondent seem to matter. It is also unclear how personal incentives affect the accuracy of response rates. We find that a lower quality interview is more likely to produce inaccurate responses, while whether the firm reports the payment of informal payments and the size of the firm are likely to reduce inaccuracies. Also, in the case of explaining differences in reported innovation between surveys, the time elapsed between interviews and interviewing the same person are strong predictors of differences, likely the result of the recall effect.

Given the prominent role that innovation plays in raising productivity and contributing to economic growth, and the need for a better understanding of firm level innovation dynamics, these findings call for a renewed effort to better measure firm-level innovation outcomes. This will require the revision of existing tools in the context of the revision of the Oslo manual, but perhaps also to test new approaches

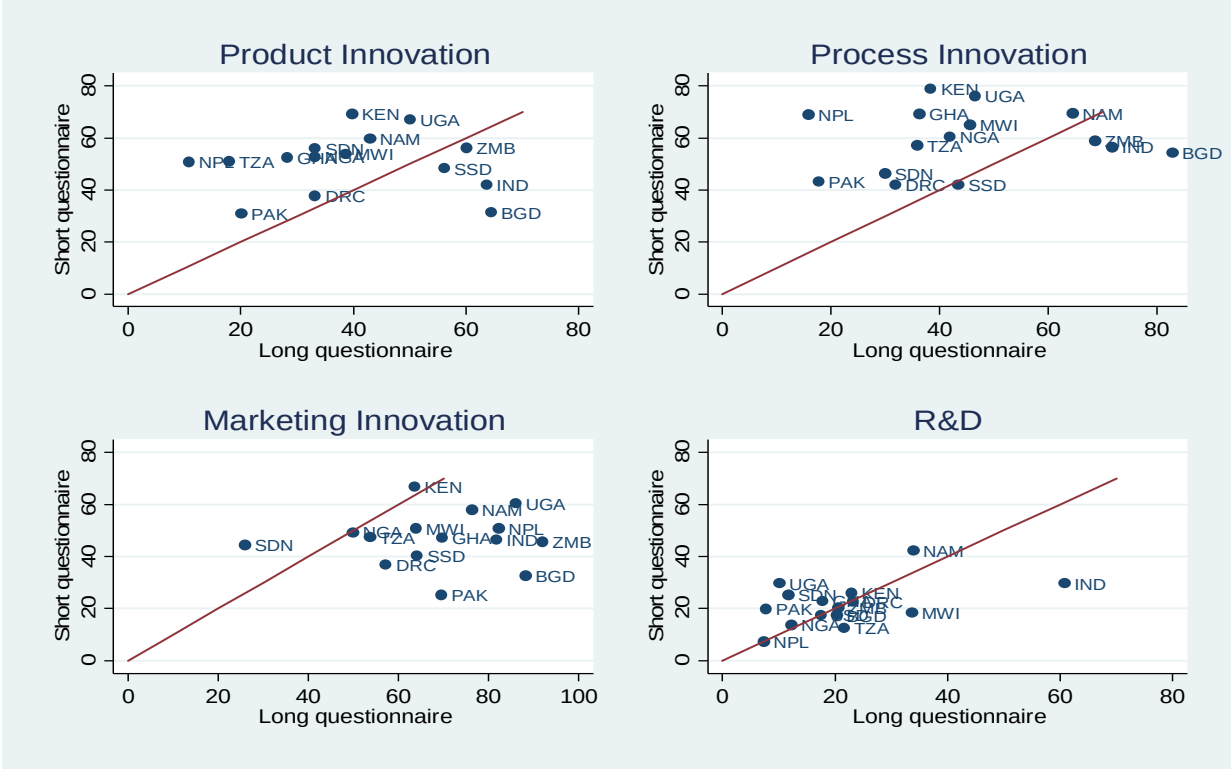
that look at firms' innovation activities more holistically, providing more clarity and less subjectivity on some of the innovation questions with the possibility of some type of verification.

***Figure 1 - Innovation rates in the main ES and innovation follow-up survey by type of innovation***



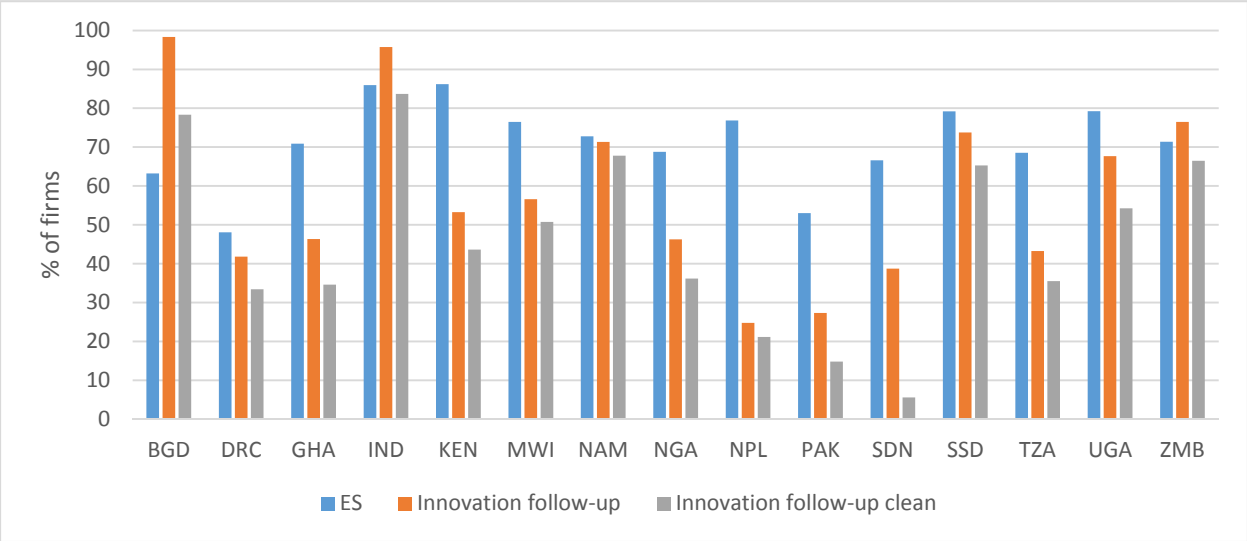

*Source: Enterprise Surveys and Innovation Enterprise Surveys*

***Figure 2 - Differences in innovation rates between short and long questionnaire***




*Source: Enterprise Surveys and Innovation Enterprise Surveys*

***Figure 3 - Innovation rates by country in ES, follow-up module and clean follow-up module***




*Source: Enterprise Surveys and Innovation Enterprise Surveys*

***Table 1: Decomposition of innovation responses in the short, long, and long-clean questionnaire***

| | no>no | yes>no | y>y>y | y>y>n | n>y>y | n>y>n | % verified | Correct in short questionnaire | Differences short and long | Long questionnaire reclassified |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| | | | | | | | (3) (4) (5) and (6) | (1) (3) and (6) | (2) (5) and (6) | (4) and (6) |
| BGD | 1.01 | 1.52 | 53.84 | 12.42 | 24.04 | 7.17 | 97.47 | 62.02 | 32.73 | 19.59 |
| DRC | 33.07 | 17.06 | 30.18 | 6.04 | 11.55 | 2.1 | 49.87 | 65.35 | 30.71 | 8.14 |
| GHA | 28.13 | 25.18 | 31.8 | 11.58 | 2.39 | 0.92 | 46.69 | 60.85 | 28.49 | 12.5 |
| IND | 2.18 | 3.12 | 71.63 | 12.96 | 8.66 | 1.46 | 94.71 | 75.27 | 13.24 | 14.42 |
| KEN | 11.19 | 33.39 | 43.67 | 7.89 | 3.3 | 0.55 | 55.41 | 55.41 | 37.24 | 8.44 |
| MWI | 12.35 | 26.34 | 42.8 | 7.00 | 9.47 | 2.06 | 61.33 | 57.21 | 37.87 | 9.06 |
| NAM | 10.37 | 32.18 | 45.21 | 4.52 | 6.12 | 1.6 | 57.45 | 57.18 | 39.9 | 6.12 |
| NGA | 23.85 | 25.46 | 37.44 | 5.18 | 5.53 | 2.53 | 50.68 | 63.82 | 33.52 | 7.71 |
| NPL | 17.23 | 45.32 | 26.38 | 4.68 | 4.47 | 1.91 | 37.44 | 45.52 | 51.7 | 6.59 |
| PAK | 52.02 | 19.73 | 10.31 | 14.2 | 1.35 | 2.39 | 28.25 | 64.72 | 23.47 | 16.59 |
| SDN | 29.1 | 32.34 | 4.23 | 29.85 | 0.75 | 3.73 | 38.56 | 37.06 | 36.82 | 33.58 |
| SSD | 14.6 | 8.5 | 63.96 | 8.87 | 3.88 | 0.18 | 76.89 | 78.74 | 12.56 | 9.05 |
| TZA | 21.67 | 33.15 | 27.96 | 7.22 | 6.85 | 3.15 | 45.18 | 52.78 | 43.15 | 10.37 |
| UGA | 9.87 | 18.39 | 45.07 | 10.31 | 12.56 | 3.81 | 71.75 | 58.75 | 34.76 | 14.12 |
| ZMB | 18.84 | 0 | 66.04 | 8.77 | 5.22 | 1.12 | 81.15 | 86 | 6.34 | 9.89 |
| **Total** | **14.39** | **15.36** | **49.09** | **10.87** | **8.01** | **2.28** | **70.25** | **65.76** | **25.65** | **13.15** |

*Source: Enterprise Surveys and Innovation Enterprise Surveys*
Note: no weights are used in the table

***Table 2: Firm characteristics by group and percentage of firms in each group***

| | | **Size** *employees* | **Age** *years* | **Exporter** *% of firms* | **Foreign** *% of firms* | **Manuf.** | **Wholesale and retail** | **Other services** |
|---|---|---|---|---|---|---|---|---|
| **Non-innovators** | no>no | 51.85 | 15.32 | 12.08% | 8.41% | 12.04% | 19.15% | 16.15% |
| | yes>no | 59.76 | 16.50 | 18.81% | 7.28% | 11.83% | 22.51% | 17.78% |
| | y>y>n | 127.67 | 18.88 | 20.09% | 5.99% | 10.24% | 12.53% | 10.82% |
| | n>y>n | 62.94 | 18.11 | 13.47% | 3.64% | 2.19% | 2.51% | 2.30% |
| **Innovators** | y>y>y | 119.98 | 17.84 | 21.61% | 8.40% | 54.52% | 37.70% | 45.65% |
| | n>y>y | 63.45 | 18.21 | 12.85% | 3.77% | 9.18% | 5.60% | 7.29% |

*Source: Enterprise Surveys and Innovation Enterprise Surveys*

***Table 3: Knowledge activities by group (percentage of firms)***

| | | **R&D** | **R&D internal** | **R&D external** | **Training** | **Equipment** | **Purchase of knowledge** |
|---|---|---|---|---|---|---|---|
| **Non-innovators** | no>no | 7.65% | 6.75% | 2.88% | 10.32% | 11.14% | 2.96% |
| | yes>no | 11.32% | 10.36% | 3.25% | 14.27% | 17.69% | 4.83% |
| | y>y>n | 32.88% | 30.72% | 9.53% | 34.67% | 43.00% | 7.25% |
| | n>y>n | 25.70% | 23.27% | 5.74% | 25.79% | 44.05% | 7.54% |
| **Innovators** | y>y>y | 41.14% | 38.42% | 9.85% | 41.80% | 60.93% | 7.70% |
| | n>y>y | 31.29% | 28.15% | 7.84% | 29.30% | 61.88% | 7.01% |

*Source: Enterprise Surveys and Innovation Enterprise Surveys*

***Table 4: Knowledge activities by group (percentage of firms)***

| | | **No show cards** | **Show cards** |
|---|---|---|---|
| **Non-innovators** | no>no | 18.14% | 18.32% |
| | yes>no | 21.32% | 23.04% |
| | y>y>n | 9.31% | 8.94% |
| | n>y>n | 1.83% | 1.85% |
| **Innovators** | y>y>y | 43.28% | 42.32% |
| | n>y>y | 6.13% | 5.54% |

*Source: Enterprise Surveys and Innovation Enterprise Surveys*

**Table 5: Comparing short and long questionnaire. Probit-Marginal effects**

Dependent variable: Different answer provided in the two questionnaires

| VARIABLES | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Firm is potential innovator Y:1 N:0 | -0.182** | -0.131 | -0.282* | -0.401** |
| | (0.089) | (0.095) | (0.154) | (0.166) |
| Manager's years of experience | -0.001 | | -0.008 | |
| | (0.004) | | (0.007) | |
| Respondent experience with firm | | 0.007 | | -0.003 |
| | | (0.006) | | (0.010) |
| Respondent is woman Y:1 N:0 | 0.166 | 0.187 | 0.286* | 0.216 |
| | (0.119) | (0.127) | (0.155) | (0.168) |
| Use of show cards Y:1 N:0 | | | 0.008 | -0.064 |
| | | | (0.125) | (0.135) |
| Time interval btw svy (days) | 0.002*** | 0.002*** | 0.002*** | 0.002*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Same respondent Y:1 N:0 | -0.231*** | -0.218** | -0.055 | -0.148 |
| | (0.082) | (0.088) | (0.147) | (0.153) |
| Nr of interviews per supervisor | 0.018*** | 0.017*** | 0.008 | 0.008 |
| | (0.004) | (0.005) | (0.013) | (0.012) |
| Enumerator's perception of truthful interview Y:1 N:0 | -0.020 | -0.026 | 0.058 | 0.169 |
| | (0.079) | (0.086) | (0.129) | (0.137) |
| Length ES interview (minutes) | -0.001 | -0.002** | -0.002 | -0.004** |
| | (0.001) | (0.001) | (0.002) | (0.002) |
| Average number of meetings with tax officials | 0.011 | 0.012 | 0.050* | 0.047 |
| | (0.010) | (0.009) | (0.029) | (0.032) |
| Firm expected to pay informal payment Y:1 N:0 | -0.154* | -0.154* | -0.285* | -0.323** |
| | (0.084) | (0.091) | (0.148) | (0.156) |
| Employment growth (%) | 0.002 | 0.002 | -0.002 | 0.001 |
| | (0.002) | (0.002) | (0.003) | (0.003) |
| Size (log) | -0.140*** | -0.128*** | -0.078 | -0.077 |
| | (0.039) | (0.043) | (0.062) | (0.066) |
| Age (log) | 0.040 | -0.034 | 0.063 | 0.039 |
| | (0.062) | (0.073) | (0.103) | (0.120) |
| Foreign ownership (25+%) | -0.024 | -0.210 | -0.284 | -0.494** |
| | (0.165) | (0.176) | (0.208) | (0.215) |
| Export (directly or indirectly) | -0.019 | 0.036 | 0.180 | 0.232 |
| | (0.111) | (0.118) | (0.163) | (0.170) |
| Constant | 0.054 | 0.241 | -0.287 | -0.067 |
| | (0.271) | (0.289) | (0.423) | (0.449) |
| Observations | 7,090 | 6,281 | 1,592 | 1,431 |

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 Country & ISIC 2-digit FE not shown

**Table 6: Comparing short and long questionnaire. Mlogit**

Dependent variable: Type of discrepancy between answers provided in the two questionnaires

| | (1a) yes_no | (1b) no_yes | (2a) yes_no | (2b) no_yes | (3a) yes_no | (3b) no_yes | (4a) yes_no | (4b) no_yes |
|---|---|---|---|---|---|---|---|---|
| Firm is potential innovator Y:1 N:0 | --1.045*** | 0.615*** | -0.956*** | 0.702*** | -0.993*** | 0.685* | -1.034*** | 0.260 |
| | (0.219) | (0.203) | (0.235) | (0.212) | (0.361) | (0.398) | (0.380) | (0.432) |
| Manager's years of experience | -0.007 | 0.010 | | | -0.026* | 0.016 | | |
| | (0.009) | (0.009) | | | (0.015) | (0.019) | | |
| Respondent experience with firm | | | -0.000 | 0.036*** | | | -0.020 | 0.047* |
| | | | (0.013) | (0.013) | | | (0.021) | (0.025) |
| Respondent is woman Y:1 N:0 | 0.237 | 0.459 | 0.264 | 0.541* | 0.715** | 0.181 | 0.517 | 0.115 |
| | (0.237) | (0.310) | (0.254) | (0.325) | (0.319) | (0.454) | (0.347) | (0.462) |
| Use of show cards Y:1 N:0 | | | | | 0.053 | -0.103 | -0.098 | -0.257 |
| | | | | | (0.239) | (0.372) | (0.260) | (0.398) |
| Time interval btw svy (days) | 0.003*** | 0.003*** | 0.004*** | 0.002** | 0.004*** | 0.003* | 0.005*** | 0.002 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) | (0.001) | (0.002) |
| Same respondent Y:1 N:0 | --0.598*** | -0.003 | -0.599*** | 0.122 | -0.342 | 0.725* | -0.467 | 0.660 |
| | (0.169) | (0.204) | (0.181) | (0.216) | (0.307) | (0.410) | (0.319) | (0.462) |
| Nr of interviews per supervisor | 0.032*** | 0.019 | 0.031*** | 0.015 | 0.021 | -0.039 | 0.017 | -0.024 |
| | (0.008) | (0.012) | (0.009) | (0.013) | (0.025) | (0.041) | (0.022) | (0.042) |
| Enumerator's perception of truthful interview Y:1 N:0 | 0.084 | -0.240 | 0.048 | -0.258 | 0.150 | -0.080 | 0.272 | 0.164 |
| | (0.159) | (0.200) | (0.167) | (0.226) | (0.251) | (0.403) | (0.260) | (0.464) |
| Length ES interview (minutes) | -0.002 | -0.001 | -0.004** | -0.001 | -0.003 | -0.004 | -0.006* | -0.006 |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.003) | (0.004) | (0.003) | (0.004) |
| Average number of meetings with tax officials | 0.022 | 0.013 | 0.021 | 0.023 | 0.110** | 0.026 | 0.091 | 0.061 |
| | (0.019) | (0.022) | (0.018) | (0.021) | (0.056) | (0.089) | (0.060) | (0.085) |
| Firm expected to pay informal payment Y:1 N:0 | -0.117 | -0.543** | -0.138 | -0.480** | -0.421 | -0.786* | -0.503 | -0.763* |
| | (0.181) | (0.223) | (0.196) | (0.239) | (0.301) | (0.440) | (0.310) | (0.456) |
| Employment growth (%) | -0.001 | 0.010* | -0.001 | 0.011* | -0.004 | -0.005 | 0.001 | -0.003 |
| | (0.004) | (0.005) | (0.004) | (0.006) | (0.006) | (0.008) | (0.006) | (0.009) |
| Size (log) | -0.072 | -0.476*** | -0.028 | -0.529*** | 0.002 | -0.552** | -0.034 | -0.450* |
| | (0.090) | (0.085) | (0.097) | (0.095) | (0.118) | (0.218) | (0.123) | (0.235) |
| Age (log) | 0.079 | -0.003 | -0.020 | -0.153 | 0.110 | 0.179 | 0.084 | -0.030 |
| | (0.130) | (0.162) | (0.149) | (0.172) | (0.208) | (0.300) | (0.230) | (0.354) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Foreign ownership (25+%) | -0.002 | -0.453 | -0.294 | -0.926** | -0.286 | -0.946 | -0.709 | -1.395* |
| | (0.334) | (0.400) | (0.342) | (0.433) | (0.419) | (0.750) | (0.432) | (0.795) |
| Export (directly or indirectly) | 0.047 | -0.117 | 0.118 | 0.041 | 0.263 | -0.029 | 0.364 | 0.277 |
| | (0.207) | (0.293) | (0.220) | (0.310) | (0.281) | (0.552) | (0.293) | (0.593) |
| Constant | -0.043 | -2.570*** | 0.224 | -2.266*** | -0.656 | -4.059*** | -0.190 | -3.725** |
| | (0.555) | (0.749) | (0.599) | (0.798) | (0.862) | (1.345) | (0.903) | (1.478) |
| Observations | 7,095 | 7,095 | 6,288 | 6,288 | 1,608 | 1,608 | 1,444 | 1,444 |

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 Country & ISIC 2-digit FE not shown

**Table 7: Probit on wrong answers in the ES. Marginal effects**

Dependent variable: Wrong answer provided in the short questionnaires

| VARIABLES | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Firm is potential innovator Y:1 N:0 | 0.294*** | 0.317*** | 0.280* | 0.230 |
| | (0.084) | (0.089) | (0.152) | (0.160) |
| Manager's years of experience | 0.003 | | -0.004 | |
| | (0.005) | | (0.008) | |
| Respondent experience with firm | | -0.002 | | 0.006 |
| | | (0.006) | | (0.012) |
| Respondent is woman Y:1 N:0 | -0.175 | -0.226* | -0.333* | -0.391* |
| | (0.117) | (0.128) | (0.185) | (0.203) |
| Use of show cards Y:1 N:0 | | | 0.109 | 0.045 |
| | | | (0.136) | (0.148) |
| Nr of interviews per supervisor | -0.006 | -0.003 | 0.005 | 0.008 |
| | (0.004) | (0.005) | (0.014) | (0.015) |
| Enumerator's perception of truthful interv Y:1 N:0 | -0.086 | -0.062 | -0.125 | -0.088 |
| | (0.074) | (0.083) | (0.144) | (0.158) |
| Length ES interview (minutes) | -0.000 | -0.000 | 0.000 | -0.000 |
| | (0.001) | (0.001) | (0.002) | (0.002) |
| Average number of meetings with tax officials | 0.004 | 0.006 | 0.003 | 0.016 |
| | (0.009) | (0.009) | (0.033) | (0.035) |
| Firm expected to pay informal payment Y:1 N:0 | -0.102 | -0.042 | -0.218 | -0.248 |
| | (0.082) | (0.089) | (0.171) | (0.184) |
| Employment growth (%) | 0.003 | 0.004 | 0.000 | 0.001 |
| | (0.002) | (0.003) | (0.004) | (0.004) |
| Size (log) | -0.119*** | -0.151*** | -0.235*** | -0.234*** |
| | (0.032) | (0.034) | (0.074) | (0.080) |
| Age (log) | 0.034 | 0.087 | 0.007 | -0.061 |
| | (0.063) | (0.078) | (0.108) | (0.137) |
| Foreign ownership (25+%) | -0.049 | -0.059 | 0.276 | 0.261 |
| | (0.156) | (0.167) | (0.245) | (0.258) |
| Export (directly or indirectly) | 0.122 | 0.162 | 0.204 | 0.226 |
| | (0.098) | (0.107) | (0.175) | (0.184) |
| Constant | -1.050*** | -1.114*** | -0.775** | -0.857** |
| | (0.249) | (0.264) | (0.385) | (0.397) |
| Observations | 7,876 | 7,005 | 1,562 | 1,405 |

Standard errors in parentheses*** p<0.01, ** p<0.05, * p<0.1 Country & ISIC 2-digit FE not shown

**Table 8: Logit on wrong answers in the ES**

| VARIABLES | (1a) yes_no | (1b) no_yes | (2a) yes_no | (2b) no_yes | (3a) yes_no | (3b) no_yes | (4a) yes_no | (4b) no_yes |
|---|---|---|---|---|---|---|---|---|
| Firm is potential innovator Y:1 N:0 | 0.537*** | 0.429* | 0.491** | 0.624*** | 0.439 | 0.640 | 0.516 | 0.157 |
| | (0.173) | (0.226) | (0.194) | (0.225) | (0.316) | (0.518) | (0.331) | (0.545) |
| Manager's years of experience | 0.002 | 0.008 | | | -0.018 | 0.028 | | |
| | (0.013) | (0.010) | | | (0.020) | (0.022) | | |
| Respondent experience with firm | | | -0.018 | 0.020 | | | -0.003 | 0.055* |
| | | | (0.014) | (0.013) | | | (0.030) | (0.030) |
| Respondent is woman Y:1 N:0 | -0.507* | -0.168 | -0.665** | -0.142 | -0.779* | -0.560 | -0.896* | -0.680 |
| | (0.274) | (0.345) | (0.306) | (0.368) | (0.455) | (0.526) | (0.481) | (0.611) |
| Use of show cards Y:1 N:0 | | | | | 0.188 | 0.128 | 0.110 | -0.142 |
| | | | | | (0.309) | (0.420) | (0.319) | (0.466) |
| Nr of interviews per supervisor | -0.025*** | 0.045*** | -0.020* | 0.043*** | 0.024 | -0.046 | 0.025 | -0.020 |
| | (0.008) | (0.014) | (0.010) | (0.014) | (0.032) | (0.049) | (0.032) | (0.048) |
| Enumerator's perception of truthful interv Y:1 N:0 | 0.054 | -0.379* | 0.116 | -0.364 | -0.197 | -0.213 | -0.176 | -0.135 |
| | (0.154) | (0.198) | (0.172) | (0.230) | (0.317) | (0.417) | (0.342) | (0.511) |
| Length ES interview (minutes) | -0.000 | -0.000 | -0.001 | -0.001 | 0.002 | -0.003 | 0.002 | -0.004 |
| | (0.002) | (0.002) | (0.002) | (0.003) | (0.004) | (0.005) | (0.004) | (0.005) |
| Average number of meetings with tax officials | 0.004 | 0.010 | 0.001 | 0.023 | 0.003 | -0.013 | 0.011 | 0.074 |
| | (0.017) | (0.020) | (0.020) | (0.018) | (0.066) | (0.099) | (0.074) | (0.093) |
| Firm expected to pay informal payment Y:1 N:0 | 0.037 | -0.529** | 0.107 | -0.379 | -0.312 | -0.814 | -0.433 | -0.711 |
| | (0.173) | (0.229) | (0.185) | (0.257) | (0.396) | (0.518) | (0.422) | (0.516) |
| Employment growth (%) | -0.000 | 0.013** | -0.001 | 0.015** | -0.002 | 0.004 | -0.002 | 0.006 |
| | (0.005) | (0.006) | (0.005) | (0.006) | (0.009) | (0.009) | (0.010) | (0.010) |
| Size (log) | -0.063 | -0.396*** | -0.082 | -0.533*** | -0.342** | -0.651*** | -0.383** | -0.574** |
| | (0.067) | (0.081) | (0.075) | (0.091) | (0.167) | (0.238) | (0.183) | (0.247) |
| Age (log) | 0.104 | 0.002 | 0.242 | 0.030 | 0.048 | -0.197 | -0.090 | -0.309 |
| | (0.138) | (0.162) | (0.178) | (0.167) | (0.263) | (0.302) | (0.339) | (0.355) |
| Foreign ownership (25+%) | 0.149 | -0.854** | 0.246 | -1.380*** | 0.745 | -0.882 | 0.784 | -1.666** |
| | (0.346) | (0.420) | (0.351) | (0.443) | (0.517) | (0.870) | (0.519) | (0.848) |
| Export (directly or indirectly) | 0.245 | 0.156 | 0.195 | 0.413 | 0.378 | 0.317 | 0.398 | 0.604 |
| | (0.198) | (0.291) | (0.226) | (0.301) | (0.371) | (0.592) | (0.383) | (0.615) |
| Constant | -2.550*** | -3.582*** | -2.756*** | -3.557*** | -1.753** | -2.714** | -1.825** | -2.773* |
| | (0.497) | (0.815) | (0.537) | (0.884) | (0.818) | (1.262) | (0.855) | (1.421) |
| Observations | 7,890 | 7,890 | 7,015 | 7,015 | 1,616 | 1,616 | 1,453 | 1,453 |

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1 ; Country & ISIC 2-digit FE not shown

**Table 9: Determinants of providing an inaccurate answer (verified sample)**

| VARIABLES | (1) wrong | (2) verified | (3) wrong | (4) verified | (5) wrong | (6) verified | (7) wrong | (8) verified |
|---|---|---|---|---|---|---|---|---|
| Firm is potential innovator Y:1 N:0 | 0.009 | | 0.055 | | -0.034 | | -0.066 | |
| | (0.086) | | (0.093) | | (0.160) | | (0.157) | |
| Manager's years of experience | 0.001 | | | | -0.009 | | | |
| | (0.006) | | | | (0.009) | | | |
| Respondent experience with firm | | | -0.000 | | | | 0.010 | |
| | | | (0.007) | | | | (0.014) | |
| Respondent is woman Y:1 N:0 | -0.185 | | -0.222 | | -0.252 | | -0.315 | |
| | (0.133) | | (0.143) | | (0.191) | | (0.211) | |
| Use of show cards Y:1 N:0 | | | | | 0.079 | | 0.013 | |
| | | | | | (0.141) | | (0.148) | |
| Nr of interviews per supervisor | 0.014*** | | 0.017*** | | 0.015 | | 0.019 | |
| | (0.005) | | (0.006) | | (0.016) | | (0.016) | |
| Enumerator's perception of truthful interv Y:1 N:0 | -0.124 | | -0.096 | | -0.219 | | -0.163 | |
| | (0.082) | | (0.092) | | (0.148) | | (0.161) | |
| Length ES interview (minutes) | -0.000 | | -0.001 | | -0.001 | | -0.002 | |
| | (0.001) | | (0.001) | | (0.002) | | (0.002) | |
| Average number of meetings with tax officials | 0.017 | | 0.018 | | 0.021 | | 0.038 | |
| | (0.013) | | (0.013) | | (0.031) | | (0.033) | |
| Firm expected to pay informal payment Y:1 N:0 | -0.251*** | | -0.195** | | -0.371** | | -0.403** | |
| | (0.086) | | (0.095) | | (0.171) | | (0.179) | |
| Employment growth (%) | 0.003 | 0.003 | 0.004 | 0.003 | -0.000 | -0.001 | 0.001 | -0.001 |
| | (0.003) | (0.002) | (0.003) | (0.002) | (0.003) | (0.003) | (0.004) | (0.003) |
| Size (log) | -0.179*** | 0.056 | -0.208*** | 0.040 | -0.238*** | 0.004 | -0.233*** | 0.025 |
| | (0.037) | (0.034) | (0.040) | (0.038) | (0.084) | (0.055) | (0.087) | (0.057) |
| Age (log) | 0.073 | -0.078 | 0.101 | 0.018 | 0.003 | -0.173** | -0.124 | -0.117 |
| | (0.072) | (0.054) | (0.085) | (0.058) | (0.124) | (0.087) | (0.154) | (0.089) |
| Foreign ownership (25+%) | 0.004 | -0.163 | -0.035 | -0.075 | 0.291 | 0.067 | 0.271 | 0.120 |
| | (0.166) | (0.139) | (0.179) | (0.148) | (0.246) | (0.182) | (0.259) | (0.188) |
| Export (directly or indirectly) | 0.164 | 0.128 | 0.226* | 0.109 | 0.248 | 0.081 | 0.281 | 0.036 |
| | (0.112) | (0.101) | (0.120) | (0.107) | (0.186) | (0.150) | (0.193) | (0.152) |
| R&D | | 0.496*** | | 0.516*** | | 0.673*** | | 0.721*** |
| | | (0.105) | | (0.110) | | (0.178) | | (0.175) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Training for innovation | | 0.343*** | | 0.348*** | | 0.337** | | 0.292* |
| | | (0.106) | | (0.109) | | (0.155) | | (0.157) |
| Equipment for innovation | | 0.884*** | | 0.911*** | | 1.055*** | | 1.063*** |
| | | (0.092) | | (0.096) | | (0.134) | | (0.134) |
| Purchase of knowledge | | 0.127 | | 0.178 | | -0.300 | | -0.230 |
| | | (0.168) | | (0.175) | | (0.283) | | (0.286) |
| Constant | -0.621* | -1.014*** | -0.633* | -1.337*** | -0.284 | -0.878*** | -0.409 | -1.195*** |
| | (0.332) | (0.213) | (0.358) | (0.224) | (0.522) | (0.313) | (0.513) | (0.315) |
| Observations | 8,550 | 8,550 | 7,854 | 7,854 | 1,868 | 1,868 | 1,747 | 1,747 |

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Country & ISIC 2-digit FE not shown

# References


Bertrand M. and Sendhil Mullainathan, 2001. "Do People Mean What They Say? Implications for Subjective Survey Data," American Economic Review, American Economic Association, vol. 91(2), pages 67-72, May.

Bogliacino, Francesco; Perani G.; Pianta M.; Supino S. 2012. “Innovation and Development. The Evidence from Innovation Surveys”. Latin American Business Review Vol. 13 Issue 3, p219-261.

Cirera Xavier, 2015. "Catching Up to the Technological Frontier?," World Bank Other Operational Studies 21684, The World Bank.

Cirera, Xavier, Leonard Sabetti and Leonardo Iacovone (2015), "Catching Up to the Technological Frontier? Understanding Firm-Level Innovation in Developing Countries, Trade and Competitiveness GP, World Bank Group, mimeo, June.

Corrado, Carol, Charles R. Hulten and Daniel E. Sichel (2005), "Measuring Capital and Technology: an Expanded Framework", in C. Corrado, J. Haltiwanger and D. Sichel (eds.) Measuring Capital in the New Economy, NBER Press.

Corrado, Carol A., Charles R. Hulten, Daniel E. Sichel (2006). "Intangible Capital and Economic Growth," NBER Working Papers 11948, National Bureau of Economic Research.

Corrado, Carol, Jonathan Haskel, Cecilia Jona-Lasinio and Massimiliano Iommi (2011), "Intangible Capital and Growth in Advanced Economies: Measurement Methods and Comparative Results", INTAN-Invest.

Crespi, Gustavo & Zuniga, Pluvia, 2012. "Innovation and Productivity: Evidence from Six Latin American Countries," World Development, Elsevier, vol. 40(2), pages 273-290.

Crespi Gustavo & Arias-Ortiz Elena & Tacsir Ezequiel & Vargas Fernando & Zuñiga Pluvia, 2014. "Innovation for economic performance: the case of Latin American firms," Eurasian Business Review, Eurasia Business and Economics Society, vol. 4(1), pages 31-50, June.

Das J., J. Hammer, C. Sánchez-Paramo. 2012. The impact of recall periods on reported morbidity and health seeking behavior. Journal of Development Economics, 98, pp. 76–88

Fernandes Ana M. & Caroline Paunov, 2015. "The Risks of Innovation: Are Innovating Firms Less Likely to Die?," The Review of Economics and Statistics, MIT Press, vol. 97(3), pages 638-653, July.

Galindo-Rueda, F. and V. Millot, 2015. “Measuring Design and its Role in Innovation”, OECD Science, Technology and Industry Working Papers, 2015/01, OECD Publishing.

Galindo-Rueda, F. and A. Van Cruysen, 2015. “Testing innovation survey concepts, definitions and questions: findings from cognitive interviews with business managers”, mimeo

Hall, B.H., 2011. Innovation and productivity. Nordic Economic Policy Review, 2, pp.167-204.

Hall, B.H., Lotti, F., and Mairesse, J., 2009. Innovation and productivity in SMEs: Empirical evidence for Italy. Small Business Economics, 33(1), pp.13-33.

Hall, B.H., Mairesse, J., and Mohnen, P., 2010. Measuring the returns to R&D. In: B.H. Hall and N. Rosenberg, eds. Handbook of the Economics of Innovation. Amsterdam: Elsevier, pp.1034-1076.

Hall, B.H., Maffioli A., 2008: "Evaluating the impact of Technology Development Funds in Emerging Economies. Evidence from Latin America". NBER wp 13835

Harkness J.A, Fons J. R. van de Vijver, Peter Ph. Mohler. 2002. Cross-Cultural Survey Methods. Wiley Series in Survey Methodology.

Harrison, Rupert Jordi Jaumandreu, Jacques Mairesse, and Bettina Peters (2008) Does Innovation Stimulate Employment? A Firm-Level Analysis Using Comparable Micro-Data from Four European Countries NBER Working Paper No. 14216 August 2008

Heckman, J. (1976), The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models in , NBER Chapters, National Bureau of Economic Research, Inc

Hoskens, M. (2015). Short vs. Long Form: Method Effects in Measuring Innovation. Paper presented at the OECD NESTI/EUROSTAT scoping workshop on the third revision of the Oslo Manual. Oslo, Norway.

Hulten, Charles R. and Janet X. Hao (2012), "The Role of Intangible Capital in the Transformation and Growth of the Chinese Economy", NBER Working Papers, No.18405, National Bureau of Economic Research (NBER), September

Jaramillo, H., Lugones, G., & Salazar, M. (2000). Manual de Bogotá - Normalización de Indicadores de Innovación Tecnológica para América Latina y el Caribe. Bogota: OEA, Colciencias, RICYT, OCyT.

Johnson, T. P., D. O'Rourke, N. Chavez, S. Sudman, R. Warneke, and L. Lacey. 1997. "Social Cognition and Responses to Survey Questions among Culturally Diverse Populations." Pp. 87-113 in Survey Measurement and Process Quality, edited by Lyberg, C., P. Biemer, M. Collin, C. Dippo, E. deLeeuw, N. Schwartz, and D. Trewin. New York: John Wiley & Sons.

Johnson, T. P. and F. J. R. van de Vijver. 2003. "Social Desirability in Cross-cultural Research." Pp. 195-204 in Cross-Cultural Survey Methods, edited by Harkness, J., F. J. R. van de Vijver, and P. P. Mohler. Hoboken: John Wiley & Sons.

Kaplan, David S. & Pathania, Vikram, 2010. "What influences firms' perceptions?," Journal of Comparative Economics, Elsevier, vol. 38(4), pages 419-431, December.

NEPAD Planning and Coordinating Agency (NPCA). 2014. African Innovation Outlook 2014, NPCA, Pretoria.

OECD (2015) *OECD Science, Technology and Industry Scoreboard*. OECD, Paris.

OECD/Eurostat, 2005. Olso Manual: Guidelines for Collecting and Interpreting Innovation Data. OECD Publishing, Paris.

Olken, Benjamin A., 2009. "Corruption perceptions vs. corruption reality," Journal of Public Economics, Elsevier, vol. 93(7-8), pages 950-964, August.

Paunov, Caroline, 2016. "Corruption's asymmetric impacts on firm innovation," Journal of Development Economics, Elsevier, vol. 118(C), pages 216-231.

Paunov, Caroline & Rollo, Valentina, 2016. "Has the Internet Fostered Inclusive Innovation in the Developing World?," World Development, Elsevier, vol. 78(C), pages 587-609.

Raffo J., S. Lhuillery and L. Miotti (2008): "Northern and Southern innovativity: a comparison across European and Latin American countries", The European Journal of Development 27 Research, 20(2): 219-239

Razafindrakoto, M and Roubaud, François 2010. ” Are International Databases on Corruption Reliable? A Comparison of Expert Opinion Surveys and Household Surveys in Sub-Saharan Africa” World Development Vol. 38, No. 8, pp. 1057–1069.

Rozkrut D.A.. 2015. “Are micro enterprises different in their innovative behavior?” On-line Journal of Current Economic and Social Topics 2015

Schumpeter, J. 1942. Capitalism, Socialism, and Democracy. New York: Harper & Bros.

Schwarz N. 1999. Self-reports: How the questions shape the answers. American Psychologist 54 (2), 93

Turriago, A. (2003): "Some aspects of technological innovation and activities in Colombia: comparison with three Latin American countries". Cuadernos de administración de la Universidad Javeriana

UNESCO. 2015. Summary Report of the 2013 UIS Innovation Data Collection. Information paper N. 24. UNESCO Institute for Statistics

Wilhelmsen L. 2014. Assessing a combined survey strategy and the impact of response rate on the measurement of innovation activity in Norway. The Statistics Newsletter for the extended OECD statistical network

Zuniga, Pluvia & Crespi, Gustavo, 2013. "Innovation strategies and employment in Latin American firms," Structural Change and Economic Dynamics, Elsevier, vol. 24(C), pages 1-17.

# Appendix 1. **Table A1 Explaining differences in measuring innovation – main variables**

| | | Variable's name | Variable description | Type of correlation with existence of discrepancies between the two surveys |
|---|---|---|---|---|
| **Cognitive Problems** | | | *Better knowledge can avoid confusion, mitigate cognitive problems and reduce discrepancies* | |
| | | Use of show cards | Additional details on innovation provided during the short questionnaire by using explanatory show cards | The more details are provided during the interview the less the likelihood of reported errors or discrepancies due to cognitive problems (-) |
| | | Potential | Familiarity with the concept of innovation (firm have abandon or on-going innovation activities) | The more firms are familiar with innovation the less the likelihood of reported errors or discrepancies due to cognitive problems (-) |
| | | Manager education /Respond's experience | Manager education/ ES Respondent's knowledge about the firm (year of experience with the firm) | The more educated (or the more experienced) the manager (the respondent) the less the likelihood of reported errors or discrepancies due to cognitive problems (-) |
| | | Respondent is a woman | Respondent to the ES survey (short questionnaire) is a woman | Female and male respondents may be characterized by a different response behavior. Direction of correlations may go in both directions |
| **Context and Framing** | **Interview's characteristics** | | *Better quality survey can mitigate the impact of cognitive problems* | |
| | | Supervisor | # interviews per supervisor during the short questionnaire | The fewer the interviews per supervisor the better the quality of the survey the less the likelihood of reported errors or discrepancies (-) |
| | | Same | Same respondent in both interviews | Same respondent is less likely to provide different answers due same understanding of innovation or for reputational concerns (-) |
| | | Truthful | Quality of the interview (short questionnaire) from enumerator perception | The better the quality of the survey the more respondent is likely to provide accurate answers the less the probability of discrepancies (-) |
| | **Context/Survey fatigue** | | *Survey fatigue can encourage respondents to lie or to reply no in the follow-up* | |
| | | Length of the ES | Length of the ES survey (short questionnaire) in minutes | The longer the length of the first interview the higher the probability of survey fatigue and the higher the probability to answer no to the second interview |
| | | Business environment | Burdensome business environment (number of visits with tax officials) | The more firms are exposed to burdensome business environment the more likely are to experience survey fatigue the less likely are to provide accurate answers (+) |
| | | Corruption | Firms exposure to corruption (firm expected to give gifts to get things done) | The more firms are exposed to corruption the less likely are to provide accurate answers (+) |
| | | Employment growth | Firm growth in employment in the last three years | The higher the growth the higher the likelihood to provide accurate answers to show clients and investors a good performance the less discrepancies (-);or the higher the incentive to lie to protect intellectual property (+) |
| | **Firm's characteristics** | | *Firms' characteristics may impact firms' incentives when answering to surveys* | |
| | | Size | Firms' size (number of employees) | The bigger the firm, the higher the likelihood to provide accurate information due to reputation risk, the less the discrepancies (-); or the higher the incentive to lie to protect intellectual property (+) |
| | | Age | Firms' years of operations | The older the firm (the less dynamic) the less the knowledge about innovation, the higher the report errors (+); the older the firm the higher the likelihood to provide accurate information due to reputation risk, the less the discrepancies (-); |
| | | Foreign | Firm has foreign ownership (at least 25% of ownership) | The more firms are exposed to foreign technology the better the knowledge of innovation the more accurate the answers (-); or 2) or the higher the incentive to lie to protect intellectual property (+) |
| | | Trade | Firm engage in international trade (at least 1% of sales) | The more firms are exposed to international trade the better knowledge on innovation the less the discrepancies (-); or the higher the incentive to lie to protect intellectual property (+) |
| **Recall Period** | | Interval | Time elapsed between ES and Innovation module | The longer the time between the two interviews the less the probability of anchoring the answer and the higher the probability to forget, the higher the probability of discrepancies (+) |

***Table A2 –Sample size by country and sector***

| | | **Innovation section in the ES** *"Short questionnaire"* | | | **Innovation follow up module** *"Long questionnaire"* | | |
|---|---|---|---|---|---|---|---|
| | Country | Manuf. | Services | Total | Manuf. | Services | Total |
| Sub-Saharan Africa | DRC | 241 | 288 | 529 | 183 | 202 | 385 |
| | Ghana | 377 | 343 | 720 | 284 | 265 | 549 |
| | Kenya | 414 | 367 | 781 | 282 | 267 | 549 |
| | Malawi | 197 | 326 | 523 | 81 | 169 | 250 |
| | Namibia | 181 | 399 | 580 | 126 | 253 | 379 |
| | Nigeria | 1,425 | 1,248 | 2,673 | 446 | 456 | 902 |
| | South Sudan | 89 | 649 | 738 | 79 | 464 | 543 |
| | Sudan | 85 | 577 | 662 | 67 | 345 | 412 |
| | Tanzania | 440 | 373 | 813 | 272 | 271 | 543 |
| | Uganda | 378 | 384 | 762 | 206 | 243 | 449 |
| | Zambia | 364 | 356 | 720 | 265 | 275 | 540 |
| | ***TOTAL AFR*** | ***4,191*** | ***5,310*** | ***9,501*** | ***2,291*** | ***3,210*** | ***5,501*** |
| South Asia | Bangladesh | 1,179 | 263 | 1,442 | 853 | 137 | 990 |
| | India | 7,163 | 2,118 | 9,281 | 2,690 | 802 | 3,492 |
| | Nepal | 242 | 240 | 482 | 235 | 236 | 471 |
| | Pakistan | 1,086 | 161 | 1,247 | 591 | 105 | 696 |
| | ***TOTAL SAR*** | ***9,670*** | ***2,782*** | ***12,452*** | ***4,369*** | ***1,280*** | ***5,649*** |
| | ***TOTAL*** | ***13,861*** | ***8,092*** | ***21,953*** | ***6,660*** | ***4,490*** | ***11,150*** |

*Source: Enterprise Surveys and Innovation Enterprise Surveys*

***Table A3 – Innovation rates in ES and follow-up module by country and type of innovation***

| | Product | | Process | | Marketing | | R&D | |
|---|---|---|---|---|---|---|---|---|
| | ES | Follow-up | ES | Follow-up | ES | Follow-up | ES | Follow-up |
| BGD | 31.43 | 64.50 | 54.22 | 82.85 | 32.58 | 88.25 | 17.30 | 17.46 |
| DRC | 37.86 | 33.22 | 41.85 | 31.91 | 36.79 | 55.55 | 22.63 | 22.09 |
| GHA | 52.50 | 28.26 | 69.22 | 36.20 | 47.03 | 69.07 | 22.92 | 16.32 |
| IND | 41.91 | 63.73 | 56.33 | 71.66 | 46.40 | 81.13 | 29.65 | 58.22 |
| KEN | 69.27 | 39.80 | 78.74 | 37.98 | 66.84 | 63.26 | 25.91 | 22.06 |
| MWI | 53.70 | 38.54 | 64.61 | 45.27 | 50.80 | 63.90 | 18.38 | 28.27 |
| NAM | 59.57 | 42.98 | 68.82 | 64.48 | 58.02 | 75.31 | 42.23 | 25.66 |
| NGA | 52.75 | 33.14 | 59.59 | 41.80 | 49.22 | 49.55 | 13.80 | 12.04 |
| NPL | 50.80 | 10.75 | 68.88 | 15.85 | 50.71 | 82.23 | 7.26 | 6.82 |
| PAK | 30.80 | 20.09 | 42.59 | 17.74 | 25.36 | 68.41 | 19.78 | 7.09 |
| SDN | 55.85 | 33.22 | 45.76 | 29.67 | 44.55 | 23.61 | 25.21 | 9.96 |
| SSD | 48.30 | 56.13 | 41.89 | 43.30 | 40.22 | 63.12 | 17.52 | 16.48 |
| TZA | 51.01 | 17.89 | 56.78 | 35.62 | 47.49 | 53.47 | 12.58 | 20.14 |
| UGA | 67.09 | 50.11 | 74.95 | 46.48 | 60.32 | 86.08 | 29.74 | 9.43 |
| ZMB | 56.09 | 60.12 | 58.73 | 68.61 | 45.76 | 92.04 | 20.39 | 19.06 |

*Source: Enterprise Surveys and Innovation Enterprise Surveys*

# Appendix 2

# Re-classification of innovation

Based on the description of the product and process innovations, these were re-classified in order to clean errors in their attribution.

The re-classification was possible for 70% of the sample of firms that answered to both the innovation section in the ES and the Innovation follow-up module. For 30% of firms it was possible to verify both product and process innovation and for 39% either product or process innovation.

| **Innovation verified** | Freq. | Percent | Cum |
|---|---|---|---|
| Both product and process innovation | 3,495 | 31.35 | 31.35 |
| Only product innovation | 1,866 | 16.74 | 48.08 |
| Only process innovation | 2,428 | 21.78 | 69.86 |
| Neither product nor process innovation | 3,361 | 30.14 | 100 |
| Total | 11,150 | 100 | |

In some cases there was not enough information to attribute an innovation, and in other cases there was confusion between product and process innovations, and between process and marketing innovations. Below are some examples of how the re-classification was implemented.

**Delivery**

- improved delivery process (additional, superior vehicles) = process innovation
- introduce delivery as new offering (not core business) = product innovation
- introduce delivery -direct sales (same core business) = marketing
- delivery method for service sector (restaurants) = process innovation
- expansion of delivery, such as additional trucks (without specifying improvements) = not innovation
- expansion of delivery to new areas or across country = marketing
- an upgrade in delivery vehicle = process innovation

**Distinction between introducing new types of product as product innovation or marketing**

- food: new recipe but not necessarily any improvement = marketing
- garments: new line, new design = product innovation (under assumption that there is product differentiation, quality improvements)
- wholesaler starts to offer new product or new range of product = product innovation

**Other**

- creation of online store = process (services)/marketing (Manufacturing)
- wholesaler opens own retailer store = product innovation
- introduced warranty = marketing
- product innovation same description as main line of business = not innovation
- Process innovation leading to product innovation = classify as both product and process
- New brand, type, design without specifying specific attribute changes = marketing
- Training of employees, improving outcomes = process