Impact of Marketing Analytics 1.0 on Entrepreneur and Firm Performance:

Field Experiment Evidence from Rwanda

Stephen J. Anderson
Pradeep Chintagunta
Rupali Kaul
Naufel J. Vilcassim*

October 30, 2021

Manuscript under development. Please do not circulate.

*University of Texas, Austin, steve@sjanderson.info; University of Chicago Booth School of Business, pradeep.chintagunta@chicagobooth.edu ; Stanford University, krupali@stanford.edu; London School of Economics, n.vilcassim@lse.ac.uk This research was supported by grants from: the University of Chicago Polsky Center for Entrepreneurship, The London School of Economics, the Stanford Graduate School of Business, the Stanford Institute for Innovation in Developing Economies (SEED) and Stanford King Center on Global Development.

Impact of Marketing Analytics 1.0 on Entrepreneur and Firm Performance:

Field Experiment Evidence from Rwanda

ABSTRACT

This paper studies the impact of analytics technology on the performance of small firms and their entrepreneur-managers. We provide insights on three unaddressed research questions: (1) What is the effect of analytics technology on firm sales and profits? (2) What is the mechanism through which this effect changes a business? (3) How does manager interaction with such technology affect their individual aptitude? Using a randomized controlled field experiment with 550 firm-owners in Rwanda, we find a positive and significant causal effect of analytics technology on firm sales (a 45% increase) and profits (a 36% increase). In terms of mechanism, we show that analytics technology leads to firms engaging in more data-driven activities that are product-related (direct effect) and also that spill over into accounting-related activities (indirect effect). In addition, using objective ability and psychological tests, we find novel evidence that interaction with analytics technology improves an individual manager's analytical abilities and financial decision making. These results have important implications for policy-makers interested in greater financial inclusion – both growing small firms in developing economies and also improving the abilities of individual entrepreneurs in these contexts.

Keywords: marketing analytics, small firm growth, entrepreneurship, emerging markets, randomized controlled field experiment

INTRODUCTION

"Information is the oil of the 21st century, and analytics is the combustion engine", said Peter Sondergaard[1] at a Gartner Symposium in 2011. The total data created worldwide is increasing exponentially and is forecast to reach 59 zettabytes in 2020 (Statista 2020). While data is an essential raw material for organizations, it means little if not coupled with the right anlaytics. As the importance of data has grown, analytics, more broadly and marketing analytics, in particular, has become one of the hot trends in the corporate world, providing the promise of increased growth and better decision making. Top-tier consulting companies, analytics firms and academic institutions have been advocating for its use (Bhandari, Singer & Scheer 2014, City *et al.* 2020 , Ransbotham and Kiron 2018). The success of firms such as Harrah's Entertainment (Loveman 2003) and Courtyard by Marriott (Wind *et al.* 1989), that successfully leveraged data and analytics, continue to be cited frequently, to promote the benefits of analytics.

Marketing analytics is the practice of measuring and analyzing customer and market data to enhance the effectiveness of marketing decisions (Lilien 2011). It can range from the use of simple statistics and data visualization techniques to the more complex use of big data and sophisticated machine learning models. The existing managerial literature suggests that organizations have already lived through two phases of analytics and are now in the phase of Analytics 3.0 (Davenport 2013). Analytics 1.0 started off in the mid-1950s with the advancement in computing technologies, which made data storage and computation possible for companies far more quickly than the unassisted human mind ever could. For the first time data on sales and customers started being recorded, aggregated and analyzed by businesses. It helped managers go beyond intuition and use data from their own firm for decision making; such as grocery stores tracking and analyzing trends in their sales data. By the mid-2000s with the advent of internet-based organizations, the era of Analytics 2.0 began. Firms started the use of big data from transactional systems internal and external to the firm, including more complex data forms like click-stream data over the internet, audio, video, image or sensor data. To complement the massive amounts of data being collected by firms, more advanced analytics and data processing tools like machine learning, artificial intelligence, distributed data-storage/processing etc. started being used. E.g. Unilever has started digitizing store checks

[1]Former Executive Vice President, Gartner Research & Advisory

by using image recognition to test for ideal share of shelf and displays for its products at retail stores . Finally, by the 2010s, the era of Analytics 3.0 began - firms started using powerful methods for data-gathering and analysis to come up with intelligent products and features for customers in addition to using analytics for supporting the operations of the firm. E.g. Amazon using clickstream user data to come up with highly targeted ads as well as product recommendations for its customers.

Even though one may think that most organizations are now in the Analytics 3.0 era, depending on their level of maturity in analytics use, firms today can be mapped onto any one of these three phases described above. In fact, some of our earlier work and pilot surveys conducted in developing countries (Uganda, Nigeria and Rwanda) show that many small firms are still stuck in the pre-analytics era (Analytics 0.0), where no data recording happens and even if they do record data, none of it is analyzed or used for any business decision making; i.e., they are largely for account-keeping purposes. Things are not much better for the small firms in developed countries either, when it comes to adoption and use of analytics in their businesses. Only about 30 % of the US small businesses[2] use analytics in sales, products or marketing areas (Sisense, 2020). Given (i) the current state of small businesses everywhere, (ii) the relevance for many firm owners and managers unexposed to any form of formal data collection or analytics, and (iii) the need to traverse Analytics 1.0 before proceeding to the higher levels of adoption, our focus in this paper is on studying the role of analytics in its most basic form on the performance of small businesses. Interestingly, a survey conducted by McKinsey & Company of 1,000 companies with more than $1 billion in revenue, shows that most senior executives feel that their company has failed to embed analytics into all areas of the organization (Bisson *et al.* 2018). Hence, there may even be sub-functions or certain geographies within bigger firms, which may be lagging behind the rest of the organization in adoption of analytics and hence, for them too our study could potentially be relevant.

Our focus on small businesses is driven by two key considerations. The first is that it is widely believed that such businesses are the lifeblood of economies (Ardic & Saltane 2011, Beck *et al.* 2005). Further, research in this area is quite limited within the marketing literature. While the advent of platform businesses has clearly magnified the ability to look at small business (e.g., independent hotels, restaurants, etc.) outcomes (see e.g., Luca 2016) the focus of that research continues to be on the larger player, i.e., the platform itself. Further, as organizations grow larger, the need

[2]Small businesses here indicates firms which have 50 or less employees

for analytics expands with the firms more likely to be in stages 2.0 or 3.0. Studying the impact of Analytics 2.0 and higher might require the firm hiring a team of ”data-analysts” dedicated to this task or delegating some of the analytics responsibilities to employees in the Marketing/Sales organization. This would likely hinder the ability to study the impact of analytics on how managers think, their ability or way of approaching the business, etc. A simple Analytics 1.0 intervention, on the other hand, can be managed and implemented fully by the firm owner. Hence, we can obtain insights into possible mechanisms by monitoring changes in the owner's decision making across various aspects of the business.

The results of introducing analytics to firms are not clear ex-ante. A few papers based on surveys, observational data and case studies show a positive correlation between analytics and firm performance (Nair *et al.* 2017, Germann *et al.* 2013 and Brynjolfsson *et al.* 2011). These studies, though informative in providing some early evidence on the success of analytics, do not provide causal evidence of its effects on firm performance, nor do they provide any insights on the mechanism through which analytics operates within the firm. Further, they are based on experiences of larger corporations, which would likely be using Analytics 2.0 or 3.0 systems. Some studies also point to continuing skepticism regarding the impact of marketing analytics. The CMO Survey (Moorman 2020), conducted with 265 top marketers at for-profit U.S. companies, shows that the use of marketing analytics in business decisions continues to slow and its contribution to company performance is quite low, showing no real gains over time. There are other studies which suggest that there is a possibility that the impact of analytics on firm performance may be marginal at best, due to its abstraction from reality and analysis paralysis (Peters and Waterman 1982). Some academic papers also make a case for the use of the non-analytic form of reasoning which includes intuition or gut-feeling, as an essential mode of reasoning when dealing with complex decisions (Barnard 1936). These studies highlight mechanisms which could lead to a negative impact on the firms which switch to a more data and analytics-based approach to decision making. Hence, one could argue for a negative or a null effect of analytics on firm performance, as well.

The quasi-experiment by Berman & Israeli 2021 is closest in spirit to our work. They identify the effect of analytics on firms utilizing staggered adoption of analytics and panel data. However, unlike our study, they don't look at the impact of marketing analytics on an individual manager and the change in the manager's performance due to interaction with the technology, rather it studies its impact at the firm level only. Next, their dependent variable is firm revenue, but the impact on overall benefit to the the firm as a result of adopting analytics, in terms of firm profitability, can't be concluded from their study. Our field experiment, involving hundereds of firms in which the technology intervention is at the level of the individual manager running the business, provides us a clean causal way to analyze the impact of marketing analytics at an individual manager level. Thus, we are able to provide insights on the changes in actions as well as aptitude of the individual manager as a result of marketing analytics adoption. We also study the impact of the analytics intervention on firm profitability along with firm sales, thereby allowing us to conduct a cost benefit analysis of analytics adoption.

In this paper we focus on three research questions. i) What is the causal impact of marketing analytics 1.0 on firm performance, i.e. firm profits and sales? ii) What is the effect marketing analytics 1.0 on *marketing* decisions/outcomes (direct effects) and *financial or operations* decisions/outcomes (indirect or spill-over effects)? iii) What is the effect of marketing analytics 1.0 on the individual entrepreneur i.e. the effect on the numerical or problem-solving ability of the individual firm owners and on their preferences regarding use of data in decision making? We implement a Randomized Controlled Field Experiment (RCFE) with entrepreneurs in Rwanda to identify the causal impact of Marketing Analytics 1.0 on firm performance (sales and profits). Our sample consisted of 550 small firms operating in the greater Kigali regions (Kigali, Muhanga and Muanze). The firms were randomized into three experimental groups –treatment, placebo and control. The treatment group had 250 firms, who receive the Market Analytics (MA) application designed by us and smartphones with internet access. The MA application helped the entrepreneurs in recording and analyzing basic business information. In addition, analysts were assigned to train them on using the app and understanding the analytics that the app displayed. The Placebo group had 50 firms, who receive just the smartphone and internet access. This formed a counterfactual group which would help with ruling out alternative explanations of the effect such as access to a smartphone or internet. The final set of 250 firms formed the control group which receives no intervention.

Our central finding is that implementing Marketing Analytics 1.0 leads to significant improvement in firm sales (45%) as well as profits (36%). The underlying route through which the effect operates is the improvement in manager aptitude to use data for problem solving, imparted by the intervention to the treated entrepreneurs. In addition, adoption of Marketing Analytics leads to entrepreneurs adopting better marketing practices and as a spillover, also adopting better accounting and financial record-keeping practices. Together, our results establish the impact of the Analytics 1.0 intervention while shedding light on the mechanisms by which the effects are manifested.

The findings from our study can be relevant for both managers/entrepreneurs as well as policy makers. As noted earlier, a lot of firms are in the Analytics 0.0 phase and our study shows that these firms can perform better by adopting Analytics 1.0. Additionally, our study also provides a percentage improvement that small firms experience as a result of adoption of Analytics 1.0, with a cost benefit analysis done for the firms in our sample. Firms, especially small firms, have constrained resources, hence they are faced with decisions on investing those resources judiciously so as to maximize their benefit. The cost benefit analysis would help them make this decision on whether to invest in Marketing Analytics or in some alternate activity. Further, policymakers including both governments and non-government bodies care about small firms becoming more efficient and effective and we see that exposure to analytics 1.0 in our sample, enabled firms to adopt better decisions and actions across marketing and accounting, which should boost their efficiency. Therefore, Analytics 1.0 and its impact on SMEs would be relevant for policymakers and managers alike.

We aim to make the following three key contributions to the academic literature in marketing. First, we provide causal evidence to show that Analytics 1.0 can indeed lead to an increase in sales as well as profitability for small firms. It is quite interesting to see that by looking at very simple analytics on the basic data about their own business, the entrepreneurs are able to improve the performance of their firms. Second, we also provide evidence on the mechanism through which the effect operates. We find that not only does the exposure to Marketing Analytics 1.0 lead to changes in entrepreneur's actions and decisions in marketing, it also leads to indirect effects through spill-overs to financial record-keeping , which leads to the entrepreneurs tracking financial records like income statement, preparation of a budget and keeping track of their inventory more closely. Finally, we find novel evidence that exposure to Analytics 1.0 leads to a change in the aptitude of the managers interacting with the technology and they are able to use data better for problem solving. To the best of our knowledge, this is the first study that shows a change at the psychological level for the individual as a result of exposure to analytics. Of note is that we are using objective tests to measure the changes in entrepreneur ability instead of only self-reported data, which should make our results more robust.

The rest of the paper is organized as follows. First we provide a brief literature review. Next, we describe the research design and the empirical specification, then we present the intervention description and checks. Subsequently we discuss the analysis and results. And lastly, we conclude with a discussion on the implications of our work.

LITERATURE OVERVIEW

As noted earlier, Analytics 1.0 could be a great way to initiate firms into performing systematic computations on their own firm's data, which could generate meaningful insights for their business. This could be relevant for firms in both developing as well as developed countries. Mckenzie & Woodruff 2015 show that more than half of the firms surveyed in their study, do not keep basic records such as daily sales of products or use those records to analyze the increase/decrease in their product sales. Their sample is representative of enterprises in low- and middle-income countries (such as Sri Lanka, Chile, Kenya, Mexico etc.). In the absence of any studies which provide us with an idea of the status of analytics usage by small firms in developing countries, this paper gives us a sense of the status of data recording which could be considered as a precursor to analytics adoption.

The activities that are measured in the existing literature are mostly financial practices such as income statements etc. Next, the firms do not use a formal technology for data recording and analytics. Importantly, these studies do not exogenously shift the use of analytics for some firms to study the impact on the firms. As the data recording itself is not wide-spread in most small firms, we can imagine the primitive state of analytics use that these firms in low-income countries find themselves in. This is not specific to low-income countries, as SMEs in more developed countries also face similar challenges. Hill & Scott 2004 conduct a qualitative study with 11 small firms in Ireland and show that even though these firms record basic sales and customer information, they did not connect the disparate business data they collected and did not analyze the data in any meaningful way. Sadok & Lesca 2009 conduct a similar study with 20 SMEs in France, they find that there is no formal process for data acquisition or storage. While, some enterprise members collected field data, it was mostly memorized in the mind of individuals who collected it and remained in an abstract form. Given the state of data collection, it was not possible for the data to be used in any analytics for decision making. These examples provide us some insights into the crude way in which data is used in many small firms even in countries with more evolved economies. As a result, we believe our study will be relevant for all the managers or entrepreneurs who are a part of such firms, and still in the early stages of analytics adoption.

Impact of Marketing Analytics 1.0 on firm sales and profits

While analytics and data science are gaining widespread popularity in organizations, there is not much academic literature that provides robust empirical evidence of the impact it has on firms. There are quite a few case studies that show the successful implementation of marketing analytics tools at some organizations, lead to a positive impact for the firm. Wixom *et al.* 2013 show how GUESS?, Inc., a clothing and retail company, used business analytics in order to identify the right product placement in stores. Bajari *et al.* 2019, study the impact of more data on the demand forecasting system for Amazon and find more data improves demand forecasts over time, though with diminishing marginal utility. Davenport and Harris (2007) even present a case study which highlights how at times using marketing analytics may end up revealing so much about the customers that it might lead to violation of customer trust. Even though we cannot use these studies to claim a robust evidence of a positive impact of analytics on firm performance, these studies are useful to further develop our understanding of the use of analytics in organizations. Nair *et al.* 2017 use a randomized trial at a single firm, to show an increase in the theoretical spending of consumers as a result of the firm using an improved customer targeting tool. While the study isolates a causal impact of one part of Analytics 3.0 (a sophisticated targeting technology) on the firm, they are unable to provide insights on the individual interaction of a manager with a complete system of analytics.

Moving on from the single firm case study based approach, Germann *et al.* 2013, use data from a survey run with 212 senior executives of Fortune 1000 firms, to study the impact of the use of marketing analytics on the performance of the firms. They complement the performance data from the survey response with firm-wise actual income and asset information. Their study finds that marketing analytics is positively associated with firm revenue. Further they also highlight several moderating factors such as support from top management team, analytical skills of employees and capturing appropriate data, which are essential to ensure effectiveness of marketing analytics. The paper however does not consider the impact of the self-selection bias of the Fortune 1000 firms that chose to adopt marketing analytics, on the performance metric of the firms. One could argue the presence of several unobservable variables that could be causing an effect on firm performance. The authors acknowledge this gap and only claim a correlation between firm performance and marketing analytics without commenting on causality. Further, this study also does not provide insights on the interaction of the individual managers with the analytics technology. Another study conducted along similar lines, is by Brynjolfsson *et al.* 2011, which provides evidence on a positive connection between data driven decision making and firm productivity, using data from 179 large publicly traded firms. While they do use an instrument variable approach to take care of reverse causality and potential endogeneity, they still are unable to find a causal relationship between firm profits and use of data driven decision making, even though the impact on sales seems to be significant.

In intent, the closest paper to our study is Berman & Israeli 2021, although the nature of the intervention, the context and the identification strategy are all quite different. They use a a quasi experiment that identifies effect of analytics on firms utilizing staggered adoption of analytics and panel data. They also show mechanism on actions taken by the online retailers across advertising, pricing and adoption of other technologies. However, similar to the other papers listed earlier, their study too, doesn't look at the impact of marketing analytics on an individual manager and the change in the manager's performance due to interaction with the technology, rather it studies its impact at the firm level only. Next, their dependent variable is firm revenue, but the impact on overall benefit to the the firm as a result of adopting analytics, in terms of firm profitability, can't be concluded from their study. Since one of the key identification techniques they use is staggered adoption of the analytics technology, their could be some self-selection issues, even though the authors present various results using matching as well as use of instrument variables to show robustness. Our field experiment, involving hundereds of firms in which the technology intervention is at the level of the individual manager running the business, provides us a clean causal way to analyze the impact of marketing analytics at an individual manager level. Thus, we are able to provide insights on the changes in actions as well as aptitude of the individual manager as a result of marketing analytics adoption. We also study the impact of the analytics intervention on firm profitability along with firm sales, thereby allowing us to conduct a cost benefit analysis of analytics adoption.

The literature also provides some evidence on possible negative effects of adoption of analytics by firms. Research shows that for corporate decisions in an uncertain environment, both intuition as well as analysis are used, with intuition overpowering analysis in most cases (Huang & Pearce 2015). Due to limited availability of time, managers worry that marketing analytics would slow them down in return for marginal or no improvement in performance (Harari 1996). Researchers have also made a case for intuition, interpersonal interaction and judgmental decision making, stating that unlike scientists, managers do not always have the time for ordered rational analysis (Simon 1987 ,Barnard 1936). Some other concerns include abstraction from reality and analysis paralysis (Peters and Waterman 1982). In the absence of a study which provides a return on investment assessment for the adoption of analytics, or conclusively shows a positive causal impact on the sales and profits of firms, it could be possible that the adoption of analytics leads to a negative impact or no impact on the firm's performance.

Based on these studies discussed we hypothesise that:

Hypothesis 1 *Marketing analytics 1.0 increases the sales and profits of the firms.*

We think it is important to note here that almost all the studies that have been conducted in the field of analytics deal with more established firms working on Analytics 2.0 or 3.0 systems. Since our aim is to study Analytics 1.0, we provide initial causal evidence of what happens to firms when for the first time, they shift to a data and analytics-based approach by using basic data from their own business and simple analytics techniques. As we have noted earlier, a large proportion of firms both in the developing as well as high-income economies, still fall in this bucket and hence these findings would be relevant for them.

Impact of Marketing Analytics on the entrepreneur's aptitude

To the best of our knowledge, within the marketing literature, there is no existing paper that studies the impact of exposure to marketing analytics or any other technology adoption on the aptitude of the managers who use it. But, due to the unique data-set that we generate in our study, we can test for this interaction of the individual manager with the analytics technology. Our sample primarily consists of firms with single decision makers who are provided with access to marketing analytics tools. The experiment has exogenous adoption of the analytics technology by individual entrepreneurs, as a result we can study the changes at a psychological level for these entrepreneurs who are exposed to the analytics application. It is important to note that, in the absence of an experiment it would be difficult to test if analytics usage leads to a change in individual's aptitude due to the intrinsic reverse causality in this context.

The rest of the literature in the broader economics area is also quite scanty, and we don't see any study evaluating the impact of a technology such as analytics on an individual manager's performance. We find some tangentially related studies on education, where there is evidence of teaching school curriculum leading to students becoming better in math. Dillon et al. (2017) show that pre-schoolers receiving math training through interactive games, leads to them having higher symbolic math skills. Banerjee et al. (2007) show remedial education and computer-assisted learning program can improve math and language scores of the students. While improvement in math skills of students is not the same as a change in aptitude of a manager, this does signal that usage of analytics could have a positive effect on improving aptitude of managers.

Behavioral studies report that low numeracy could lead to biased decision making - susceptible to moods and context factors (Reyna *et al.* 2009). Studies show that individuals with better numeracy end up choosing normatively better options (Pachur & Galesic 2013). Paulos 1988 shows that low ability to deal with probabilities and small likelihoods of large outcomes, results in misinformed personal decisions and government policies, as well as an increased reliance on pseudo-science. Most of the literature in this area comes from psychology, and it does point to the importance of aptitude in the day to day decision making by individuals. We feel this should be of benefit in a commercial setting too, for running day to day business operations in a more effective manner. This brings us to our next hypothesis:

Hypothesis 2 *Usage of marketing analytics leads to improvement in entrepreneurs aptitude which improves their ability to use data for problem-solving, thereby acting as one mechanism for the increase in sales and profits for the firm.*

Impact of Marketing Analytics on Marketing activities of the firm and spill-over effects on non-Marketing activities

We study the actions and decisions of the entrepreneurs in our sample, in order to uncover the underlying mechanism of change. One of the more obvious changes that might happen to firms as a result of adoption of marketing analytics is that they might start adopting better business practices. The response of the manager as a result of being exposed to the outputs from the marketing analytics tool is a key aspect, which could determine the effectiveness of analytics adoption. To the best of our knowledge, there isn't any existing literature that evaluates these changes in actions or decisions at the manager level as a result of exposure to marketing analytics. We study managerial actions directly related to technology i.e. Marketing based and indirectly related to the technology i.e. financial recording and accounting based.

Berman & Israeli 2021 show that as a result of analytics exposure, firms integrate better customer prospecting technologies to their websites which in-turn help in attracting more paid visitors to their websites. Bajari *et al.* 2019 state that by using analytics for demand forecasting, Amazon.com is able to link its purchase ordering system to the forecasting system that helps in efficient inventory management minimizing cases of stock-outs or over-stocks. While further adoption of related technologies might be an obvious consequence of adoption of analytics, we feel that it could also lead to changes in day to day marketing and non-marketing decisions of the entrepreneurs too. Marketing analytics might provide helpful insights to the managers which nudges them to adopt better practices across finance, operations etc. Thus our last hypothesis is :

Hypothesis 3 *Usage of marketing analytics leads to -*

A. *firms changing their actions within marketing-related business functions (direct-actions), thereby acting as a mechanism for the increase in sales and profits for the firm.*

B. *spill-over effects as a result of firms changing their actions within non-marketing business functions such as accounting and finance (indirect-actions), thereby acting as another mechanism for the increase in sales and profits for the firm.*

RESEARCH DESIGN

To empirically study the impact of marketing analytics 1.0 on firm performance would be difficult using observational data. This is due to – (1) difficulty in getting data for analytics 1.0, i.e. when firms start to record, aggregate and analyze data from their business for the first time. It may not be possible to identify firms moving from a clean slate where there is no analytics use, to a scenario where they start using data and analytics for their business. In most cases, the firms for which data are publicly available are already exposed to analytics in some form and may be using it to arrive at decisions related to their businesses. Hence, they will not be suitable candidates for analysis. (2) Even if we are able to identify such a data-set of firms adopting marketing analytics 1.0, we would still have endogeneity concerns, since firms with certain internal characteristics or external situations may be more likely to opt for analytics 1.0. If these internal or external variables are also correlated with firm performance, it would lead to incorrect causal interpretations of the impact of analytics 1.0 on firm sales or profits. (3) We may also face the problem of reverse causality, wherein changes in firm performance may influence the analytics adoption decision of the firms rather than vice-versa. Given these identification concerns, running an RCFE addresses all these issues and provides us with clean causal inferences of the impact of marketing analytics 1.0 on firm sales and profits.

Empirical Context: Rwanda

The choice of Rwanda for this study was made since we could easily identify small firms which were completely unexposed to any form of data analytics and hence, as discussed earlier, this provided us with an opportunity to measure the impact of Marketing Analytics 1.0 on these firms. We could exogenously shift the exposure to Marketing Analytics for a random subset of the firms and measure the changes, if any, amongst the firms. Further, in the recent years, Rwanda has had a strong government push towards a technology revolution to make it the next Silicon Valley. The government has built initiatives to expand internet connectivity as well as access to technology with the goal of transforming their society into a highly digitized country . This push came in handy to us, from a logistics standpoint, as internet connectivity was stable for our recruited sample. Further, the government's zeal for improving small business performance helped us to get the necessary permissions from the government for running the study.

Recruitment and Sample

We obtained our sample of entrepreneurs through door to door recruitment of firms. We hired an on-ground team of 20 research assistants who canvassed the streets of Kigali, Muhanga and Musanze to recruit firms for our study. They approached a total of about 15,000 firm owners of which 3,140 completed the recruitment survey. Then, we used a Growth Index to identify the top 1,500 firms that were growth oriented and hence most suitable to be a part of this study. The Growth

Index was calculated using the data collected in the recruitment survey, and is a sum of 10 key components evaluated on a total of 100 possible points. These 10 components covered metrics which helped us assess the firm's and entrepreneur's characteristics[3]. Using the Growth Index for selecting our sample ensures that we include firms which are operational, will be willing to participate and adopt our intervention and have a proper physical structure (which increases the possibility of being able to locate these firms a few months after launching the intervention). We excluded the firms which were at subsistence levels, unwilling to participate in our intervention and without a fixed physical structure.

Baseline Data Collection

We reached out to these 1,500 shortlisted firms, and were able to find 954 firms during the baseline, that agreed to be a part of the project. However, we experienced technical delays in the application development and ran another baseline about 3 years later when we were finally ready to launch the intervention[4]. The detailed timeline for the project is presented in Appendix Figure A1. This final baseline (which is what we are using in this paper) consisted of 550 firms, the remaining firms from the previous baseline were either not found or they refused to be a part of the study. The baseline survey contained questions on the business and entrepreneur background and business performance. The firms in our sample on average had 0.5 full-time employees, an asset value of 2,640 USD, last month sales of 707 USD and last month profits of 113 USD. About 45% of the entrepreneurs in our sample were female; average age was 37 years and only 13.5% of the entrepreneurs had attended a prior business education or training program. The firm owners were not highly educated either and more than 87% of them never went to college. Due to these characteristics of the sample, we ensured that the manipulation we introduce to increase the analytics exposure of the firms, was

[3]The 10 components are - firm endowments (which is the start-up capital), if the firm has an established business location, number of paid employees of the business, business effectiveness (e.g. record keeping, separate finances), experimentation (new activities or innovations), entrepreneur education (traditional schooling and business training), previous work experience of the entrepreneur (salaried job or a job at a larger/corporate environment), entrepreneur's exposure (if the entrepreneur has visited or lived in other countries), external research (considering others' perspectives) and finally enumerator's evaluation (i.e. enumerator assessment of the entrepreneur)

[4]We ran one more baseline with 1044 firms, approximately a year after completing the first baseline. We tried to contact the top 2000 firms as per our original recruitment survey, so as to be able to get a large enough sample size, of which we were able to find 1044 firms who agreed to be a part of the study. This baseline was conducted in anticipation of launching the intervention, however we experienced a major bottleneck in the application development which resulted in further delays.

simple and easy to understand. We will describe this in more detail in the subsequent sections.

Randomization and Balance Checks

After we conducted the baseline survey (May - June 2019) for the 550 firms in our sample, we randomized them into the three experimental groups – Treatment, Placebo and Control. Group 1 was the treatment (n=250) which received the Marketing Analytics 1.0 intervention (smartphone and app), Group 2 was the placebo (n=50) which received just the smartphone with the internet connection and Group 3 was control (n=250) which received no experimental manipulation. The Placebo group was incorporated in the experiment so we could rule out alternative explanations linked to exposure to the internet or smartphone use. We conducted balance checks at the beginning of the intervention to test the randomization. These are reported in Appendix Table A1. The data from the baseline survey was used for conducting the balance checks on the sample. We used 39 variables in total, which include firm performance variables, business characteristics and entrepreneur characteristics. The firm performance variables include the sales and profits of the firm. The business characteristics include number of employees (full-time and part-time), assets (total value, tangible assets, working capital), number of business partners, products (number of products, product margins) and customers (number of regular and new customers, basket size of purchase). Finally, the entrepreneur characteristics include education, age, prior ownership of a business and orientation towards use of data. As can be seen from Appendix Table A1, the sample is balanced when tested across these variables at the baseline.

Endline Data Collection

Post randomization, we started the marketing analytics intervention in July 2019 which continued for seven months up to January 2020. Finally, seven months after having started the intervention, we conducted the endline survey from February to March 2020, which concluded right before the country-wide lock-down due to COVID-19 began in Rwanda. Refer Figure A1 for the detailed timeline of the experiment. The endline survey included most of the variables measured at the baseline such as entrepreneur and business characteristics. The focus in the endline survey was also to collect data on the mechanism which included marketing activities, spill-over activities to the finance and operations functions of the firm and the entrepreneur level ability. We also repeated the balance checks at the endline for the firms that could be tracked post attrition (n=527) and again the sample was balanced (results presented in Appendix - Table B1). This makes us confident that the randomization worked well and attrition did not cause the sample to become unbalanced. However, in the results section of the paper, we conduct robustness checks, by running LASSO regressions for the main effects as well as the mechanisms by including all the variables on which we conducted the balance checks as covariates.

Attrition Checks and Firm Survival

As mentioned earlier, we were able to conduct the endline survey for 527 of the 550 firms of our sample, thus experiencing an attrition of 23 firms which we were unable to contact or locate. These attrition numbers are quite low and non-systematic between the experimental groups. The same can be seen in Appendix Table A2. Of these 527 firms, 33 firms shut-down due to various reasons. Table A2 shows that firm survival is about 4.3% higher for the treatment group as compared to the control. The results that we report in the later sections are all ITT (Intention to Treat) effects. But, we also report the main-effects conditional on survival and the results are robust.

INTERVENTION DESCRIPTION AND CHECKS

The intervention ran for 7 months. Our main objective was to introduce marketing analytics into firms which have been unexposed to the use of data or analytics in the past. Thus we needed a manipulation that could help these small Rwandan firms in – i) capturing the business data which would act as the input for marketing analytics, ii) performing the basic analytics and presenting simple tables and data visuals which could be insightful for the business and iii) understanding the analytics output that gets generated as an end product of this manipulation. For this purpose, we designed a "Market Manager" mobile phone application (app) which would enable the firms in both recording the information and performing the analytics on the recorded data which could then be presented to the firms for use. The app enhances the firm owner's ability to access, track and take action on marketing analytics. It was important to ensure that the application is user-friendly and easy to understand, especially since majority of firm-owners had not gone to college and many of them had never used a smartphone before. The first time exposure of analytics for the firms in our sample, along with the relatively simple statistics and visual representation of their business data, ensured that we delivered an intervention in the Analytics 1.0 sphere.

We provided an Android smartphone with a fast and stable internet connection to the treatment and placebo entrepreneurs. On each of the smartphones provided to the treatment firms, the Market Manager application was installed and the entrepreneur was made to log-in using a unique ID and password. The entrepreneur used the mobile application for both data entry and to look at the analytics results which the app generated each month based on the data that was entered by them. We also hired a team of 18 data analysts and an on-ground research manager to oversee the team of data analysts (and conduct surprise visits to keep a check on the data analysts). Each data analyst was in-charge of managing about 15 treatment firms. The job of the data analysts was two-fold. First, they helped the entrepreneur in understanding the data entry process in the app both at the daily and weekly level. Second, they helped the entrepreneurs interpret the analytics output of the application linking it to the data that they had entered into the application. Critically however, these analysts provided *no* consulting or inputs into how the owners could use the data for decision-making. See Appendix Figure A2 for a schematic of the intervention set-up. As mentioned earlier, most of the entrepreneurs in our sample were not technologically savvy and many of them were using a smartphone for the first time, the initial support of the data analysts was required to ensure data entry compliance by the firm owner.

For the first two months of the intervention (July to August 2019), the data analysts visited the entrepreneurs on a weekly basis to familiarize them with the smartphone, the application use, what each data entry question in the app meant and most importantly, how they could enter data into the app. After two months, the data analysts visited each treatment entrepreneur on a monthly basis to explain the analytics output from the application. These monthly visits continued till January 2020. Overall, the feedback that we collected from the treatment group at the end of the experiment shows that the firms found the use of Analytics 1.0 useful. On a scale of 1 to 7 (1 corresponds to ”Strongly Disagree” and 7 to ”Strongly Agree”), on average, the entrepreneurs rated their satisfaction with the intervention as 6.28. On whether Analytics 1.0 represented a good value for their time, they provided an average rating of 5.55. About 76% of the entrepreneurs were willing to pay for the intervention, if it had not been offered for free.

The app focused on four key aspects of the business for recording data as well as conducting analytics – daily sales, products, customers and marketing activities. Below we describe in detail each of these. The reason we focused on these topics which covered revenue, product margin analysis, customer segment analysis, promotions and customer relationship, was that these form a key part of the syllabus of the core marketing course at most business schools. We looked at the syllabus for the introductory marketing core course at the top 10 B-schools as per Financial Times rankings and these topics formed a part of each of their curriculum. These topics, embedded in our analytics app, are central to Marketing and considered as the basic tools which managers can equip themselves with in order to improve the decisions they make regarding any business.

Market Manager Application Data Entry

The entrepreneurs were expected to enter four sets of data into the app periodically. First, the daily sales amount of the products they sold (in Rwandan Francs) was to be entered by the entrepreneur each day after they closed their shop. Next, at a bi-weekly frequency, the products, marketing activities and customer data were entered. The first and third weeks of the month were dedicated to entry of product and marketing activity information and the second and fourth weeks were dedicated to capturing customer information. The product data captured included sales, price and cost for each of the top three selling products for the firms. The marketing activity data captured included the number of times the firm had changed price of some product in the past week to promote sales, number of times promotional activities were undertaken by the firms (e.g. flyers, discount coupons, surprise gifts etc.), and the number of times the firms' owners interacted with their customers (to understand their needs, to build a closer relationship with their customers, to enquire about customer satisfaction and to enquire about customer appetite for a new or improved product). The customer data was captured for loyal, irregular and new customer segments and for each segment it included the number of customers, basket size of purchase, business clients and the number of customers from outside their neighborhood. Some snapshots of the app and the data recording process can be seen in Appendix Figure A4.

We experienced a ramp-up period in compliance to data entry for the initial two months of the launch of the intervention. The data analysts worked closely with the entrepreneurs to explain the usage of the smartphone, the application and the data entry process. The app also generated a pop-up notification twice a day (5pm and 9pm Rwanda time) prompting the entrepreneur to enter the daily or weekly data in case it was still pending. As can be seen in Appendix Figure A3, we started with a data entry compliance rate of 68 % in July 2019, but from September 2019, we experienced high data entry compliance rates of more than 90 % till January 2020. The high data entry rate shows that the entrepreneurs were not missing out on capturing the daily and weekly data entry schedules. This ensured that the analytics output that was generated at the end of each month was based on data that closely tracked the business, thereby making it more relevant.

Market Manager Application Analytics Output

Each month, based on the data captured for the month, the application generated a standard set of predefined analytics which included simple statistics and visualization of the data entered. The output was available to the entrepreneurs in the mobile phone app. In addition, every month, the data analysts delivered a hard copy of the same analytics report that was available in the app to the firm owner (about 35 pages in total). The data analyst then sat with the entrepreneur and explained how each table and graph in the report was developed. For these entrepreneurs who were not used to looking at any data related to their business, the analytics could be overwhelming to understand on their own and hence the data analysts provided them with explanations.

The analytics output had four sections to it, which mapped well to the data entry sections. The first part was the sales section which presented a summary of the daily, weekly and monthly sales trends along with the average sales numbers. The month on month sales graphs were presented to help the entrepreneurs get a picture of their firm performance and how it increased or decreased across time. The second section of the report was on products (top 3 selling products of the entrepreneur). For each of the top three products, it provided the over time (weekly, monthly) volume sold, average margin, sales value, and profit value. The pie chart and bar chart of the volume sold, and profits earned for each of the products, could help the entrepreneur in clearly viewing which product sells more in quantity and which helps them generate more profits. The third section of the report was on customers (loyal, irregular and new customer segment). This part presented over time (weekly, monthly) customer-segment wise average basket size per customer visit, customer count and amount of sales. Here too the data visualization in terms of pie charts and bar charts, could help the entrepreneur compare the customer segment-wise performance, e.g., it enabled the firm-owner to view which segments were driving maximum revenue and by how much.

Finally, the marketing activities section of the report presented the time trend across weeks, of the marketing activities (promotions, price changes, customer contacts for post-purchase satisfaction etc.) undertaken by the firm. Appendix Figure A5 shows the images of some of the output screens from the app that have the analytics results, Appendix Figure A6 shows snapshots of the analytics output in the report which was used to print out the hard-copy. Each entrepreneur in the treatment group received 7 analytics reports in total at end of each month from July 2019 to January 2020.

Taking Actions Based on the Analytics

The mobile application as well as the analytics report supplied to the entrepreneurs as a hard copy provided a standardized list of generic actions owners could take in each section. All owners could access this list which was not customized in any way. For example, the sales section provided a list that included the question "What can you do to increase the number of units sold?" with actions such as the lowering of prices. The list was pre-coded into the app even before the launch of the intervention and did not change over time. Indeed some actions would have been inappropriate in the context of a specific firm. The list was included to make it comparable to dashboards available on similar commercial software.

To reiterate, and importantly, none of the data analysts provided any information or advice to the entrepreneurs on how to use the analytics output in their business. The data analysts were trained to specifically provide the entrepreneurs an understanding of how the data entry and analytics in the application work. In fact, they were especially told not to provide any additional advice to the entrepreneurs on how to use the analytics results.

In addition, we performed manipulation checks to test if the treatment entrepreneurs paid attention to the analytics information that the app provided them. We measured some metrics that are directly linked to the analytics topics presented to the entrepreneurs each month in the report. We wanted to check if the entrepreneurs were following those topics closely and were getting accustomed to using that information. We find that the treatment firms proved to be significantly higher in the use of overall data and analytics, daily sales information, customer information and product information (Table A3). This shows that the treatment firms were paying attention to the topics in the report, just like we intended.

ANALYSIS AND RESULTS

The empirical specification we follow is an ANCOVA to measure the ITT (Intention to Treat) effects of being offered the treatment and the placebo. We define the exact specification below for firm 'i' in the Endline follow-up survey:

$$Y_i = \beta_0 + \beta_1 Treatment_i + \beta_2 Placebo_i + x_i' \theta + \gamma Y_{i,t=0} + \varepsilon_i \tag{1}$$

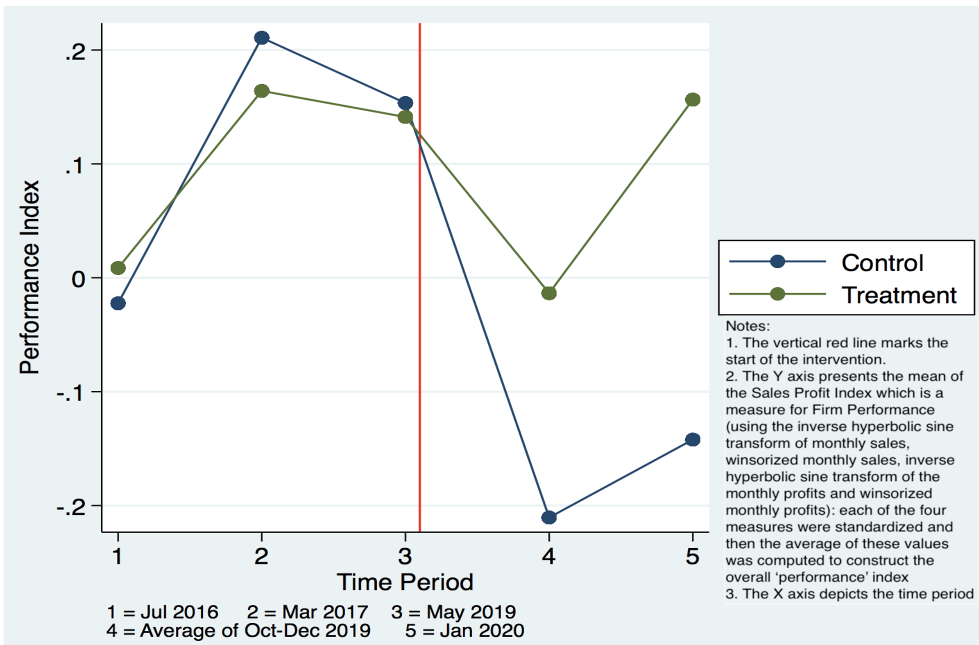
Here Y_i is the given outcome variable measured post-treatment, $Y_{i,t=0}$ is its baseline value and ε_i is the error term. β_1 (and β_2) will then provide the intent-to-treat effects, which are the effects of being assigned to the treatment (or placebo) relative to being a firm in the control. Since random assignment is at the individual firm level, robust (Eicker-Huber-White) standard errors are used. Even though our sample is balanced, as a part of the robustness checks, we use the double-lasso method of Belloni *et al.* 2014 to select the covariates to include as control, using the list of variables which are used in the balance checks. These controls should also help in correcting for any non-random attrition related to these baseline variables, even though the attrition does not seem to be systematic. We focus on intention-to-treat effects for the most part of the analysis as it provides an unbiased estimate of the impact of the treatment on firm performance (e.g., sales, profits). These ITT results represent the cleanest identification of treatment effects given they rely on an exogenous source of variation (randomization into experimental groups).

We first report the impact of marketing analytics 1.0 on firm performance measures of sales and profits. Then we move on to providing evidence on the mechanism of business level actions that include direct effects on marketing activities and the spillover effects on financial and operations activities of the business. Lastly, we will provide evidence on the changes in the numerical and reasoning ability of the entrepreneurs as well as their preference towards use of data and numbers. All the regressions, unless otherwise noted, report the intention to treat effects and thus include all firms including non-survivors (except for the attrited firms).

Impact of Marketing Analytics 1.0 on Firm Performance (Main Effect)

We first present some model free evidence by plotting the graphs of the monthly sales and profits of the firms. Looking at Figure 1 (Additional evidence from Figures A7 and A8 in Appendix), it seems that the treatment is doing much better than the control and placebo groups.

Figure 1: Model free evidence of main effects - Performance Index



It can be seen from the figures that the sales and profits for the pre-intervention period was not very different between the treatment and control groups, but post the intervention, we see an improvement in the treatment group's performance relative to the control. Next we present the regression outcomes (note that the values for sales and profits are presented in 1000 RWFs). The results for the main effects on sales and profits are presented in Table 1.

It shows the mean effect on the two measures of performance – average monthly sales and average monthly profits. The average monthly sales is the average of the last month's sales (for Jan 2020) and the typical month's sales (for the months of Oct - Dec 2019) of the firm. Similarly the average monthly profits is the average of the last month's profits (for Jan 2020) and typical month's profits (for the months of Oct - Dec 2019).

We also create an overall index – Sales Profit Index, which is an average of the standardized

IHS (inverse hyperbolic sine) transforms of the sales and profit measures of performance listed above. This index helps us in avoiding multiple hypothesis testing and captures all the relevant performance measures into a single metric (McKenzie 2017). Table 1 suggests that all of the sales as well as profits metrics are significant for the treatment firms but not the Placebo firms. The same is true for the overall Sales Profit Index as well. Looking at the last month's sales and profits, we see that the Marketing Analytics 1.0 leads to about 45 % improvement in sales and 36 % improvement in profits for the treatment firms. Since the Placebo group does not seem to show any significant change in sales or profits, we can safely rule out alternative explanations such as the effects of the Treatment group being driven by the exposure to a smartphone or the access to internet.

Also, looking at Figure 2, we see that the effects on performance are all across the distribution of the firms and are not being driven by a few outliers.

This is encouraging to see, since this means the intervention had an impact across different firm sizes. Lastly, we run a set of robustness checks. We ran the same main effects analysis for survived firms only and the results continue to hold (See Appendix Table B2). This shows the effects are not just being driven by the extensive margin of improved chances of survival, but the intensive margin of actual firm performance is also improving for the treatment group. We ran LASSO regressions for the main effects by including the baseline entrepreneur and business characteristics. We also include another regression with all the entrepreneur and business characteristics as covariates as another robustness check and the treatment effects still hold in all these cases (Appendix Table B4).

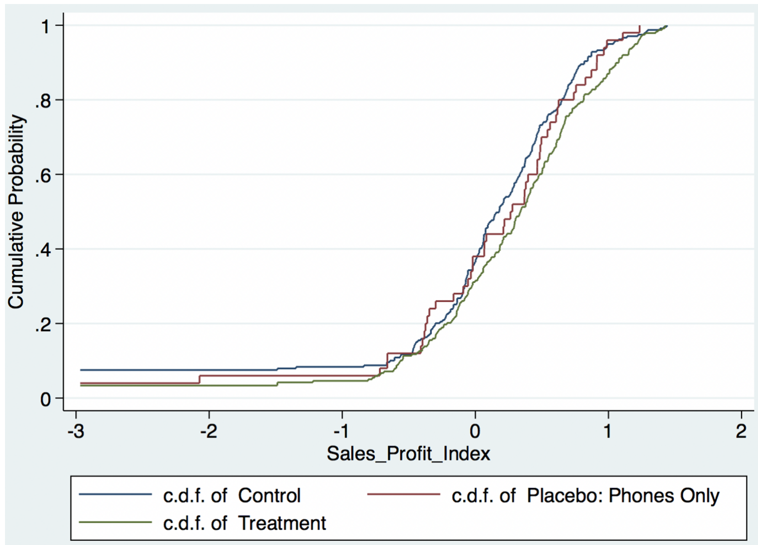
Note that we did not observe any increase in the time spent on the business by the treatment entrepreneurs. In Appendix Table B5 we see that the percentage of days in a month for which the shop is open did not change for the treatment firms across time.

Table 1: Impact of Marketing Analytics on Firm Sales and Profits

	DV: Monthly Sales		DV: Monthly Profits		Overall
	(1)	(2)	(3)	(4)	(5)
	Average	IHS	Average	IHS	Performance Index
Treatment	165.3***	0.795***	24.98***	0.621**	0.250***
	(58.60)	(0.293)	(7.939)	(0.263)	(0.0703)
Placebo	-35.20	0.255	11.88	0.475	0.0823
	(81.65)	(0.499)	(12.03)	(0.398)	(0.0965)
Mean of DV: Control	454.9	12.28	85.60	10.94	0
SD of DV: Control	706.5	3.633	79.29	3.217	0.814
Effect Size in SD: Treatment	0.234	0.219	0.315	0.193	0.307
Effect Size in %: Treatment	36.35	79.5	29.18	62.1	
Effect Size in SD: Placebo	-0.0498	0.0703	0.150	0.148	0.101
Effect Size in %: Placebo	-7.739	25.5	13.88	47.5	
Obs.	527	527	527	527	527
$\beta\_treat = \beta\_placebo$	0.0193	0.257	0.304	0.698	0.0859

Notes: This table summarizes analysis for the main effect of marketing analytics on firm performance outcomes (from baseline to endline). The results in each column represent the intention-to-treat effects (ITT) of the randomly assigned intervention to a treatment group entrepreneur. Column (1) presents a winsorized measure for Average Monthly Sales (it refers to the average sales for the last full calendar month for the firm, which for our sample was January 2020) and the sales for a typical month (average for the months Oct 2019 to Dec 2020) : estimates after winsorizing each 1% on both tails. Column (2) presents an IHS transformed measure for Average Monthly Sales: estimates after transforming the average monthly sales with the inverse hyperbolic sine function. Column (3) presents a winsorized measure for Average Monthly Profits (it refers to the average profits for the last full calendar month for the firm, which for our sample was January 2020) and the profits for a typical month (average for the months Oct 2019 to Dec 2020) : estimates after winsorizing each 1% on both tails. Column (4) presents an IHS transformed measure for Average Monthly Profits: estimates after transforming the average monthly profits using the inverse hyperbolic sine function. Column (5) presents an index measure for Firm Performance (IHS last month's sales; IHS typical sales monthly; IHS last month's profits; IHS typical profits monthly): each of the four measures were standardized and then the average of these values was computed to construct the overall 'performance' index. Column (6) presents an index measure for Firm Performance (using IHS average sales monthly; IHS average profits monthly): each of the two measures were standardized and then the average of these values was computed to construct the overall 'performance' index. All regressions include firms that failed as at endline (non-operational with zero monthly sales or profits). Any firm growth values in levels (sales; profits) are listed as Rwandan Francs (RWF) in 1000s. The regressions also include: the baseline value of the dependent variable. The IHS results are robust to the scaling of the profits/sales numbers. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1$ $p < 0.05^{**}$ $p < 0.01^{***}$

Figure 2: CDF of the Overall Performance Index



We also compare the monthly sales data from the endline survey to the sales data as reported by the treatment entrepreneurs in the analytics app. We see that the correlation between the two is 97.44%. We see that the endline survey reported last month sales tracks very well with the sales reported in the analytics application for the same period by the treatment entrepreneurs (Appendix Table B6).

We also provide additional evidence to support the main effects that we pick up, by looking at the product and customer related outcomes (Appendix Table B7) of the firms.

The product related outcomes show that the treatment group sells more of their (top 3) products and end up generating significantly higher revenue and profits from them too, as compared to the control group. The customer outcomes show that the amount of goods being purchased by customers from the treatment firms is increasing as is the referrals provided to other potential customers by the existing customers. These product and customer outcomes help us justify the increase in profits and sales of the firms that we capture in the main effects.

We present additional evidence of robustness of the main effects by performing attrition bounding analysis (Appendix Table B8) and calculating the treatment on treated effects (Appendix Table B9). We see that our treatment effects continue to hold in all scenarios of the attrition bounding analysis. Further, we see that the ATT effects (average treatment-on-treated) are higher than the ITT effects which proves that our treatment is working as we had expected. Lastly, we see that the control firms are not experiencing statistically different sales if there is a competing treatment firm nearby (Appendix Table B10). This provides evidence that the treatment firms are not performing better at the expense of control firms of our sample. Overall, this analysis presents evidence in favor of hypothesis 1.

Mechanism of the Impact of Marketing Analytics 1.0 on Firm Performance

Having established a significant increase in the main effect outcomes of the treatment firms, we now shift our focus to understand the possible mechanisms that might be leading to these changes in sales and profits for the firms.

Impact of Marketing Analytics 1.0 on the Numerical Ability of the Entrepreneurs

Hypothesis 2 proposes that one of the key mechanisms through which marketing Analytics impacts the sales and profits of the firms could be by enhancing the numerical ability of the firms. To test this mechanism, we study the difference in numerical ability between treatment and control. We use a combination of six psychological tests in order to measure the ability of the entrepreneurs to work with numerical problem solving and reasoning. The Cognitive Reflection Test, Digit Span Test, Raven's Test and Numerical Aptitude Test are strictly objective in nature; entrepreneurs were presented with a set of problems with a correct answer, and they had exactly 60 seconds to solve each problem. The Mental Math Calculation and Numerical Orientation Tests are self-report tests, where the entrepreneurs are asked about their ability and preference towards the use of numbers, data and graphs. Figure 3 presents the graphs of the entrepreneur performance on each of these six tests, across the three experimental groups. Looking at these graphs, it seems the Treatment entrepreneurs are having better analytical ability than the other groups which is encouraging. Table 2 shows the regression results for the tests that we conducted to measure the entrepreneur's numerical ability as well their preference towards use of data and analytics. We see that on each of those tests the Treatment group firms perform significantly better than the control and the placebo firms. The overall analytics ability index is also significantly higher for the Treatment group, which shows support for hypothesis 2; adoption of marketing analytics 1.0 makes the entrepreneurs more numerically able, thereby acting as one mechanism for the increase in sales and profits for the firms.

Figure 3: Model Free Evidence for Entrepreneur's Analytics Ability

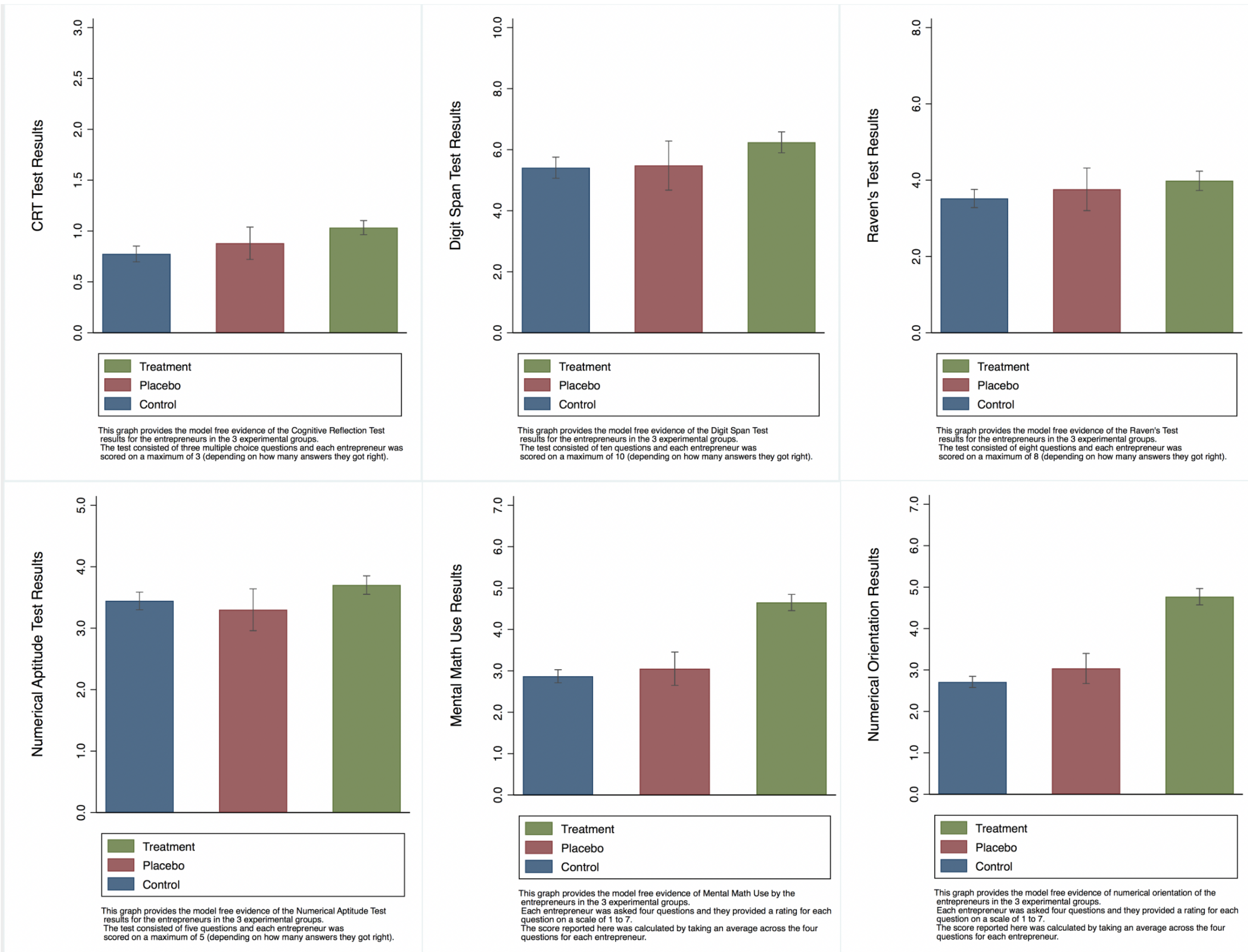


Table 2: Impact of Marketing Analytics on Entrepreneurs' Analytical Ability

	(1) CRT	(2) Digit Span Test	(3) Raven's Test	(4) Numerical Aptitude	(5) Mental Math Uses	(6) Numerical Orientation	(7) Analytical Ability Index
Treatment	0.260***	0.829***	0.464***	0.258**	1.782***	2.057***	0.763***
	(0.0531)	(0.247)	(0.177)	(0.105)	(0.129)	(0.121)	(0.0611)
Placebo	0.106	0.0700	0.241	-0.144	0.181	0.324*	0.108
	(0.0879)	(0.435)	(0.301)	(0.183)	(0.215)	(0.192)	(0.0923)
Mean of DV: Control	0.774	5.410	3.519	3.444	2.869	2.711	0
SD of DV: Control	0.614	2.716	1.878	1.121	1.254	1.053	0.595
Effect Size in SD: Treatment	0.423	0.305	0.247	0.230	1.422	1.952	1.284
Effect Size in %: Treatment	33.53	15.33	13.20	7.497	62.11	75.85	
Effect Size in SD: Placebo	0.172	0.0258	0.128	-0.128	0.144	0.307	0.182
Effect Size in %: Placebo	13.69	1.293	6.854	-4.168	6.300	11.94	
Obs.	527	527	527	527	527	527	527
$\beta\_treat = \beta\_placebo$	0.0744	0.0804	0.463	0.0303	2.17e-12	3.12e-16	2.99e-11

Notes: This table summarizes analysis for the effect of marketing analytics on entrepreneurs' analytical ability (from baseline to endline). The results in each column represent the intention-to-treat effects (ITT) of the randomly assigned intervention to a treatment group entrepreneur. Column (1) presents a measure for the Cognitive Reasoning Test score (it is measured on a scale of 0 to 3). Column (2) presents a measure for Digit Span Test score (it is on a scale of 0 - 10). Column (3) presents a measure for the Raven's Test score (it is on a scale of 0 - 8). Column (4) presents a measure of the Numerical Aptitude Score (it is on a scale of 0 - 5). Column (5) presents a measure of the Mental Math Use (it is self-reported on a scale of 1 to 7). Column (6) presents a measure of Numerical Orientation (it is self-reported on a scale of 1 to 7). Column (7) presents an index measure for Entrepreneurs' Analytical Ability. Each of the six analytical measures were standardized and then the average of these values was computed to construct the overall 'analytical ability' index. All regressions include firms that failed as at endline (non-operational with zero monthly sales or profits) as we were able to run these tests with the owners of firms that were shut-down too. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1$ $p < 0.05^{**}$ $p < 0.01^{***}$

Even though we had imposed a strict time limit on the entrepreneurs to answer each of the questions in the objective tests, we also test for the difference in interview duration across the experimental groups as a part of our robustness checks (See Appendix Table A4). But we did not find the interview duration to be significantly different across the groups, which assures us that the results that we show in this section are not caused due to a difference in the attention towards the survey across the experimental groups. We also wish to highlight the unique nature of our evidence regarding entrepreneurs' improvement in ability at a psychological level as a result of interventions aimed at improving firm sales or performance.

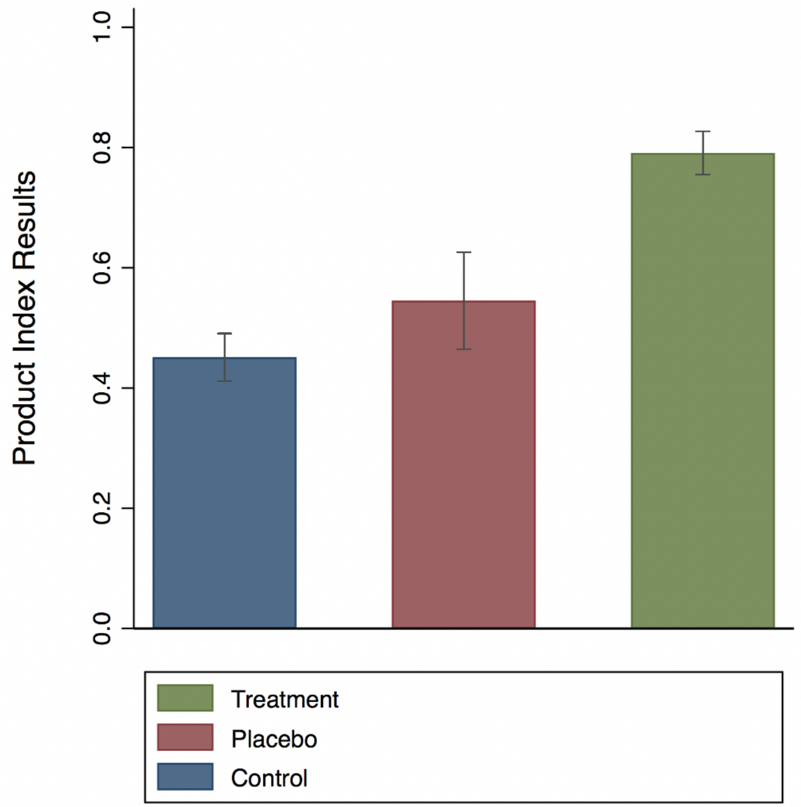
Impact of Marketing Analytics 1.0 on the Activities Performed by the Firms

Next, we look at the activities that the firm performs, as an explanation for the improvement in outcomes, i.e., do firms adopt better business practices related to business performance (Mckenzie & Woodruff 2015). We test for the differences in adoption of Product-related and Accounting-related best practices across firms. We also test for certain other activities which we do not think should be impacted as a result of analytics adoption. These other practices that we include in our analysis act as a additional robustness check as they provide divergent validity to our results.

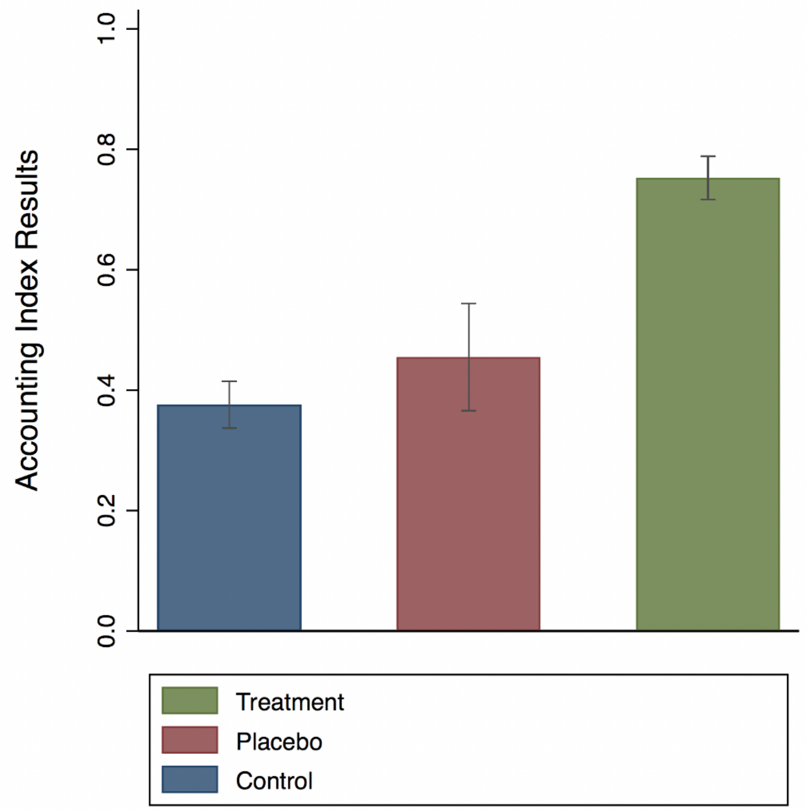
First we look at the model free evidence in Figure 4. The graphs suggest that the treatment group seems higher in adoption of best practices in business across products-related and accounting-related activities . Next we look at the regression results. Table 3 shows that Analytics 1.0 impacts the product-focused activities that the firms adopt, such as changes in prices of products, improvement/addition of products, inspecting the products for their quality etc. The overall Product Activities index is also significantly higher for the treatment group as compared to the control and the placebo.

While our Marketing Analytics intervention had a direct impact on adoption of Product related practices by the firms, we also looked at any spillover impact on other functions. Table 3 shows the impact of Analytics 1.0 on Accounting or record-keeping practices of the firms. We see that the treatment firms start recording and creating documents based on financial data such as - creating an income statement, creating a future budget and monitoring cash flow of the business. Figure 5 shows some of the data and records that the treatment firms started maintaining to aid them in their business.

Figure 4: Model Free Evidence for Business Actions



This graph provides the model free evidence of product-related activities undertaken by the entrepreneurs in the 3 experimental groups.
Each entrepreneur was asked four questions by our auditors and a yes (1) or no (0) answer was recorded for each question.
The product score reported here was calculated by taking an average across the four questions for each entrepreneur.



This graph provides the model free evidence of accounting-related activities undertaken by the entrepreneurs in the 3 experimental groups.
Each entrepreneur was asked four questions by our auditors and a yes (1) or no (0) answer was recorded for each question.
The accounting score reported here was calculated by taking an average across the four questions for each entrepreneur.

Table 3: Impact of Marketing Analytics on Business Activities

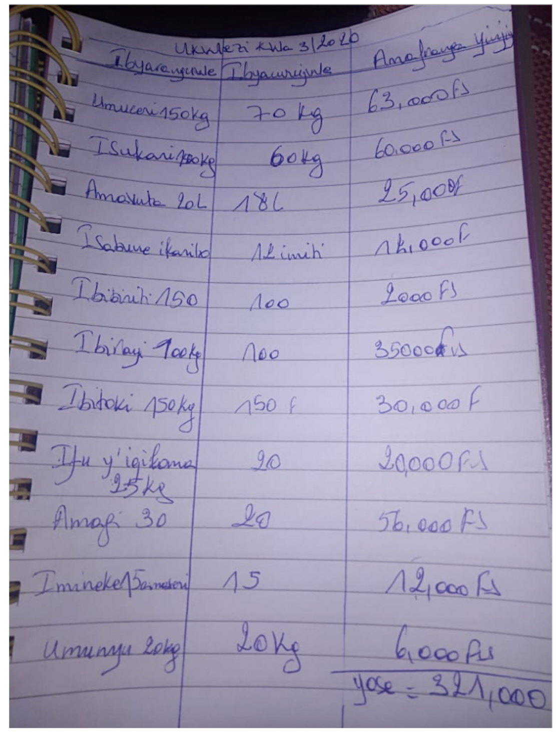
	Product Related				Accounting Related				Overall
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Change Prices	Improve/Add Products	Inspect Products	Monitor Product Inventory	Separate Business Finances	Create Income Statement	Create Future Budget	Monitor Cash Flow	Analytical Activity Index
Treatment	0.489***	0.492***	0.238***	0.141***	0.376***	0.421***	0.433***	0.276***	0.359***
	(0.0399)	(0.0396)	(0.0367)	(0.0404)	(0.0417)	(0.0398)	(0.0414)	(0.0357)	(0.0235)
Placebo	0.163**	0.168**	0.00310	0.0431	0.0983	0.129*	0.0533	0.0356	0.0866**
	(0.0766)	(0.0727)	(0.0739)	(0.0719)	(0.0777)	(0.0678)	(0.0744)	(0.0731)	(0.0411)
Mean of DV: Control	0.297	0.192	0.657	0.657	0.402	0.151	0.307	0.644	0.413
SD of DV: Control	0.458	0.395	0.476	0.476	0.491	0.358	0.462	0.480	0.270
Effect Size in SD: Treatment	1.067	1.246	0.500	0.297	0.765	1.176	0.937	0.575	1.329
Effect Size in %: Treatment	164.5	255.8	36.24	21.53	93.52	279.8	141.1	42.81	86.74
Effect Size in SD: Placebo	0.356	0.424	0.00651	0.0906	0.200	0.361	0.115	0.0743	0.321
Effect Size in %: Placebo	54.85	87.04	0.471	6.561	24.48	85.89	17.37	5.532	20.96
Obs.	527	527	527	527	527	525	526	527	527
$\beta\_treat = \beta\_placebo$	0.0000194	0.0000155	0.000858	0.161	0.000284	0.0000500	0.000000386	0.000490	4.33e-11

Notes: This table summarizes analysis for the effect of marketing analytics on activities that the business performs (from baseline to endline). The results in each column represent the intention-to-treat effects (ITT) of the randomly assigned intervention to a treatment group entrepreneur. These results are based on a Yes/No response on each of these activities, as recorded by the auditors in-person review of each of the firms in our sample. Column (1) presents a measure for changing the prices of products/services to increase total sales. Column (2) presents a measure for introducing new products/services, or improving your existing products/services, to provide more benefits to customers. Column (3) presents a measure for inspection of products to ensure they are of the highest quality. Column (4) presents a measure of monitoring stocks of supplies to prevent stock overflow or stock-outs. Column (5) presents a measure of keeping business finances separate from personal finances. Column (6) presents a measure of creating an income statement to track all the money that comes 'in' (sales) and goes 'out' (purchases, expenditures) from the business. Column (7) presents a measure for creating a budget that states how much is expected in sales and costs for a future period (e.g. next month). Column (8) presents a measure for tracking how much cash is available in the business to ensure there is enough money for daily operations to continue. Column (9) presents an index for the analytical activities performed by the business. This overall index is calculated by taking an average of each of the eight business measures. All regressions include firms that failed as at endline (non-operational with zero monthly sales or profits, and their business activity response is considered to be 'no'. All regression exclude the firms that attrited during the Endline survey round. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1$ $p < 0.05^{**}$ $p < 0.01^{***}$

Figure 5: Examples of financial records maintained by firms

Daily product-wise revenue and cost information being noted for all sales

Stock for all products being noted and tracked



This is a spillover effect of our intervention, perhaps as a result of understanding the benefit of recording and analyzing marketing related business data the treatment firms also started maintaining additional records to support them in running their firms effectively.

Appendix Table B11 shows the impact of Marketing Analytics on business activities such as production and process improvements. We do not find any significant impact on these activities, as predicted since these activities are seemingly unrelated to the exposure to marketing analytics. Hence, again we show that hypothesis 3 also holds and marketing analytics causes firms to adopt better product-related practices and has spillover effects on adoption of better accounting related practices too, which could be another mechanism of improvement of firm performance.

Mediation Analysis for the Proposed Mechanisms

Lastly, we test if the mechanisms that we captured above are indeed mediating the main effect on firm sales and profits that we had shown earlier. We run the mediation analysis for numerical ability and management practices as constructs independently mediating the sales profit index of the firms. We find that for numerical ability, the average causal mediation effect (ACME) significantly mediates the main effect and explains 87.35% of the total effect (Appendix Table B12). Similarly, for management practices, we find the ACME to be very significant (Appendix Table B13). Within the assumptions of sequential unconfoundedness, these mediation tests further provide evidence that marketing analytics led to improvement in firms' performance through the mechanisms of improvement in numerical ability of the entrepreneurs and adoption of better business practices by the firms.

CONCLUSION

Our analysis finds that Marketing Analytics 1.0 has a positive and significant impact on firm performance, as it leads to an improvement in both sales as well as profits of the firm. These results are robust to alternative specifications of the model and inclusion/exclusion of non-surviving firms. Further, we find evidence to support the hypothesis that exposure to Marketing Analytics 1.0 leads to a significant improvement in the analytical ability of the entrepreneurs. We show this improvement with the help of a battery of objective and subjective psychological measures. We rule out alternative explanations of attention, using the interview duration analysis. Next, we show evidence on Marketing Analytics 1.0 leading to an adoption of analytically driven activities by the firms, which are related to the products. We also show spillover effects on accounting and recording activities. We do not pick up any effect on other process oriented activities which were not directly related to the Marketing Analytics 1.0 exposure, thus providing discriminant validity to our results. Finally, we conduct causal mediation analysis and show that both the analytical ability of the entrepreneur and the analytical activities of the business significantly mediate the performance index (sales and profits) of the firm.

On 14 March 2020 Rwanda recorded its first case of COVID-19. About a week later, the government declared its first country-wide lockdown to contain the spread of COVID-19, becoming the first country in Africa to declare a lockdown. Borders were closed, as was inter-city travel, and all non-essential businesses were shut-down while essential businesses (medicines, groceries etc.) were allowed to open under restrictions, for about a month. Then, by mid January 2021 the country saw a renewed surge in COVID cases, thus leading to another lockdown for about 40 days, where businesses were completely shut-down (as was inter-city travel) except for essential businesses which were allowed to open with strict restrictions. The businesses started opening again from 1st March 2021, post this second lock-down. In March 2021, we conducted a follow-up qualitative survey with a set of randomly selected 14 treatment firms from our sample to check for the use of marketing analytics by the firms and to see if the treatment firms are still sticking to the learnings they received during the experiment. We had not kept in touch with these firms for about a year (since our Endline survey concluded in March 2020). As can be seen from Appendix Figure B1 , 13 of the 14 entrepreneurs that we surveyed kept some sort of data record for their business. In addition, 11 of the 14 entrepreneurs analyzed the data they recorded and used it to make business decisions. Finally, 9 of the 14 entrepreneurs claimed that the use of data analytics helped them through the COVID-19 related disruption caused to their business. Some of the benefits from the use of data analytics reported by these 9 firms, that helped develop resilience in their business to COVID related shocks, have been described in Appendix Figure B1.

We think that these results have important implications for marketing researchers, policy-makers and industry practitioners. By providing insights into how the impact of analytics takes place and by showing novel mechanisms - entrepreneur's analytical ability and firm's analytical activities. While this paper has begun to uncover how Marketing Analytics 1.0 impacts firm performance. Future research can look at studying the impact of more advanced forms of Analytics 2.0 and 3.0, which are also being adopted especially by various big companies.

For policy-makers driven to improve small firms' performance, our research provides support for the use of analytics interventions which could enable entrepreneurs improve their businesses. Further, we use a technology based solution (Mobile Application) for this purpose which is easily scalable thereby leading to lower costs and better reach if such a solution were to be implemented.

Finally, as described in the earlier sections, a number of entrepreneurs are still unsure about the merit of investing in building analytics ability in their businesses. Our study provides evidence to industry practitioners in favour of investing in analytics to improve their sales and profits.

References

1. Ardic, M. & Saltane. *Small and medium enterprises: A cross-country analysis with a new data set* (The World Bank, 2011).
2. Bajari, P., Chernozhukov, V., Hortaçsu, A. & Suzuki, J. *The impact of big data on firm performance: An empirical investigation* in *AEA Papers and Proceedings* **109** (2019), 33–37.
3. Barnard, C. I. Mind in everyday affairs. *The functions of the executive,* 301–321 (1936).
4. Beck, T., Demirguc-Kunt, A. & Levine, R. SMEs, growth, and poverty: cross-country evidence. *Journal of economic growth* **10,** 199–229 (2005).
5. Belloni, A., Chernozhukov, V. & Hansen, C. Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies* **81,** 608–650 (2014).
6. Berman, R. & Israeli, A. *The Added Value of Data-Analytics: Evidence from Online Retailers* tech. rep. (working paper, 2021).
7. Bhandari, Singer & Scheer. Using marketing analytics to drive superior growth — McKinsey & Company. *McKinsey & Company.* `http://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/using-marketing-analytics-to-drive-superior-growth` (2014).
8. Bisson, P., Hall, B., McCarthy, B. & Rifai, K. Breaking away: The secrets to scaling analytics. *McKinsey Analytics,* 1–14. `https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/breaking-away-the-secrets-to-scaling-analytics%7B%5C%%7D0Ahttps://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/breaking-away-the-secrets-to-scaling-analytics?cid=oth` (2018).
9. Brynjolfsson, E., Hitt, L. M. & Kim, H. H. Strength in numbers: How does data-driven decisionmaking affect firm performance? *Available at SSRN 1819486* (2011).
10. City, N. Y., Customer, W., Initiative, A. & Challenge, A. A. How Data Analytics Can Boost Any Organization ' s Performance, 1–6 (2020).
11. Davenport, T. H. Analytics 3.0 (2013).
12. Germann, F., Lilien, G. L. & Rangaswamy, A. Performance implications of deploying marketing analytics. *International Journal of Research in Marketing* **30,** 114–128. ISSN: 01678116. `http://dx.doi.org/10.1016/j.ijresmar.2012.10.001` (2013).
13. Harari, O. Quotations from Chairman Powell: A leadership primer. *Management Review* **85,** 34–37. ISSN: 00251895. `http://search.ebscohost.com/login.aspx?direct=true%7B%5C&%7Ddb=bth%7B%5C&%7DAN=9707300277%7B%5C&%7Dsite=ehost-live` (1996).
14. Hill, J. & Scott, T. A consideration of the roles of business intelligence and e-business in management and marketing decision making in knowledge-based and high-tech start-ups. *Qualitative Market Research: An International Journal* **7,** 48–57. ISSN: 13522752 (2004).
15. Huang, L. & Pearce, J. L. *Managing the Unknowable: The Effectiveness of Early-stage Investor Gut Feel in Entrepreneurial Investment Decisions* **4,** 634–670. ISBN: 0001839215597 (2015).

16. Lilien, G. L. Bridging the academic-practitioner divide in marketing decision models. *Journal of Marketing* **75,** 196–210. ISSN: 00222429 (2011).

17. Loveman, G. Diamonds in the data mine. *Harvard business review,* 109–113 (2003).

18. Luca, M. Reviews, reputation, and revenue: The case of Yelp. com. *Com (March 15, 2016). Harvard Business School NOM Unit Working Paper* (2016).

19. Mckenzie, D. & Woodruff, C. Nber Working Paper Series Business Practices in Small Firms in Developing Countries. `http://www.nber.org/papers/w21505` (2015).

20. Moorman. *The CMO Survey – Highlights and Insights Report, February* tech. rep. (2020). `https://cmosurvey.org/wp-content/uploads/2020/02/The%7B%5C_%7DCMO%7B%5C_%7DSurvey-Highlights-and%7B%5C_%7DInsights%7B%5C_%7DReport-Feb-2020.pdf`.

21. Nair, H. S., Misra, S., Hornbuckle IV, W. J., Mishra, R. & Acharya, A. Big data and marketing analytics in gaming: Combining empirical models and field experimentation. *Marketing Science* **36,** 699–725 (2017).

22. Pachur, T. & Galesic, M. Strategy selection in risky choice: The impact of numeracy, affect, and cross-cultural differences. *Journal of Behavioral Decision Making* **26,** 260–271 (2013).

23. Paulos, J. A. *Innumeracy: Mathematical illiteracy and its consequences* (Macmillan, 1988).

24. Peters and Waterman, R. RH (1982): In Search of Excellence: Lessons from America's Best Run Companies. *Harper \& Row (New York)* **1,** 360 (1982).

25. Ransbotham and Kiron, D. Using analytics to improve customer engagement. *MIT Sloan Management Review* (2018).

26. Reyna, V. F., Nelson, W. L., Han, P. K. & Dieckmann, N. F. How numeracy influences risk comprehension and medical decision making. *Psychological bulletin* **135,** 943 (2009).

27. Sadok, M. & Lesca, H. CAHIER DE RECHERCHE n 2009-12 E5 A business intelligence model for SMEs based on tacit A Business Intelligence Model for SMEs Based on Tacit Knowledge. *Communications* (2009).

28. Simon, H. A. Making Management Decisions: the Role of Intuition and Emotion. *Academy of Management Executive* **1,** 57–64. ISSN: 0896-3789 (1987).

29. Statista. *statistic_id871513_information-created-globally-2010-2024.pdf* 2020. `statistic%7B%5C_%7Did871513%7B%5C_%7Dinformation-created-globally-2010-2024.pdf`.

30. Wind, J., Green, P. E., Shifflet, D. & Scarbrough, M. Courtyard by Marriott : Designing a Hotel Facility with Consumer-Based Marketing Models. *Interfaces* **19,** 25–47. ISSN: 0092-2102 (1989).

31. Wixom, B. H., Yen, B. & Relich, M. Maximizing value from business analytics. *MIS Quarterly Executive* **12** (2013).

Appendix

Table A1: Balance Checks with the sample at Baseline

Variable	(1) **Control (N=250)** Mean [SE]	(2) **Placebo (N=50)** Mean [SE]	(3) **Treatment (N=250)** Mean [SE]	(4) **Total (N=550)** Mean [SE]	(5) **F-test** Joint orthogonality
Capital with which Business was started (RWFs)	3.87e+05	3.27e+05	4.74e+05	4.21e+05	1.384
	[37852.098]	[47560.813]	[54532.307]	[30572.814]	
Paid full-time employees working at the business	0.584	0.340	0.428	0.491	0.900
	[0.128]	[0.109]	[0.064]	[0.066]	
Number of Business Partners	0.088	0.260	0.084	0.102	5.633***
	[0.021]	[0.075]	[0.020]	[0.015]	
Total value of Assets of the Business (RWFs)	2.71e+06	1.68e+06	2.52e+06	2.53e+06	0.681
	[3.72e+05]	[4.20e+05]	[3.77e+05]	[2.44e+05]	
Number of different products sold at the shop	3.464	3.020	3.696	3.529	1.109
	[0.190]	[0.385]	[0.203]	[0.131]	
Number of loyal customers	30.760	24.160	19.804	25.180	0.928
	[8.273]	[4.400]	[1.455]	[3.841]	
Number of new customers	47.092	29.300	28.268	36.918	1.264
	[12.662]	[5.056]	[1.925]	[5.847]	
Basket-size of a loyal customer per visit (RWFs)	4023.362	3602.447	5557.940	4683.076	1.519
	[533.720]	[855.380]	[850.896]	[463.966]	
Basket-size of a new customer per visit (RWFs)	3362.571	6567.708	3128.327	3539.566	1.662
	[702.357]	[4149.033]	[422.540]	[522.487]	
Last month sales (30 days, RWFs)	7.08e+05	6.68e+05	7.62e+05	7.29e+05	0.193
	[63890.066]	[1.65e+05]	[89309.109]	[52045.251]	
Last month profits (30 days, RWFs)	1.28e+05	1.27e+05	1.38e+05	1.32e+05	0.238
	[7019.585]	[16961.824]	[14074.110]	[7305.860]	
Rely on ?gut-feeling? instead of ?data analytics?	2.876	3.040	2.988	2.942	0.652
	[0.079]	[0.187]	[0.081]	[0.054]	
Organizing customers into segments	4.008	3.840	4.048	4.011	1.293
	[0.054]	[0.108]	[0.053]	[0.036]	
Calculating profit margins	4.384	4.440	4.500	4.442	1.967
	[0.044]	[0.091]	[0.039]	[0.028]	
Age of entrepreneur	36.955	35.304	37.341	36.980	1.071
	[0.533]	[0.999]	[0.622]	[0.383]	

*The value displayed for F-tests are the F-statistics. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.*

Figure A1: Experimental Timeline

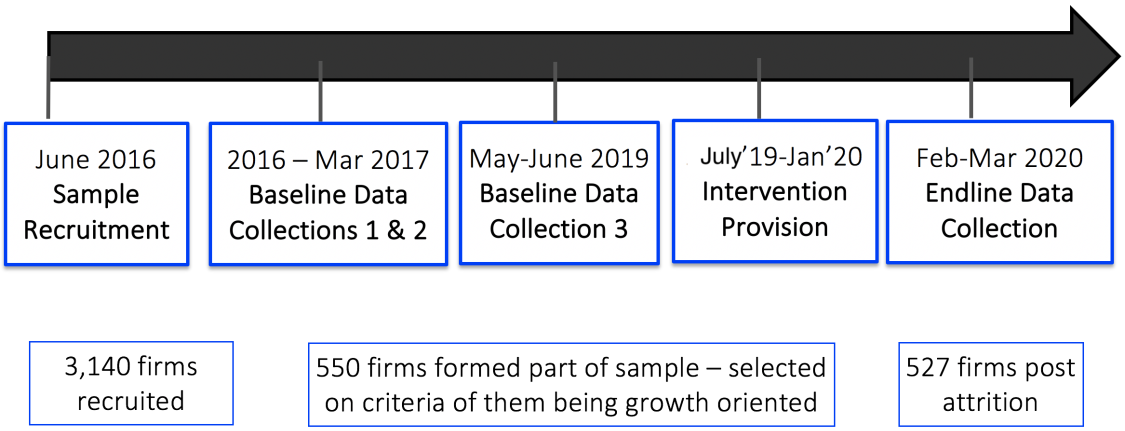


Figure A2: Intervention Set-up

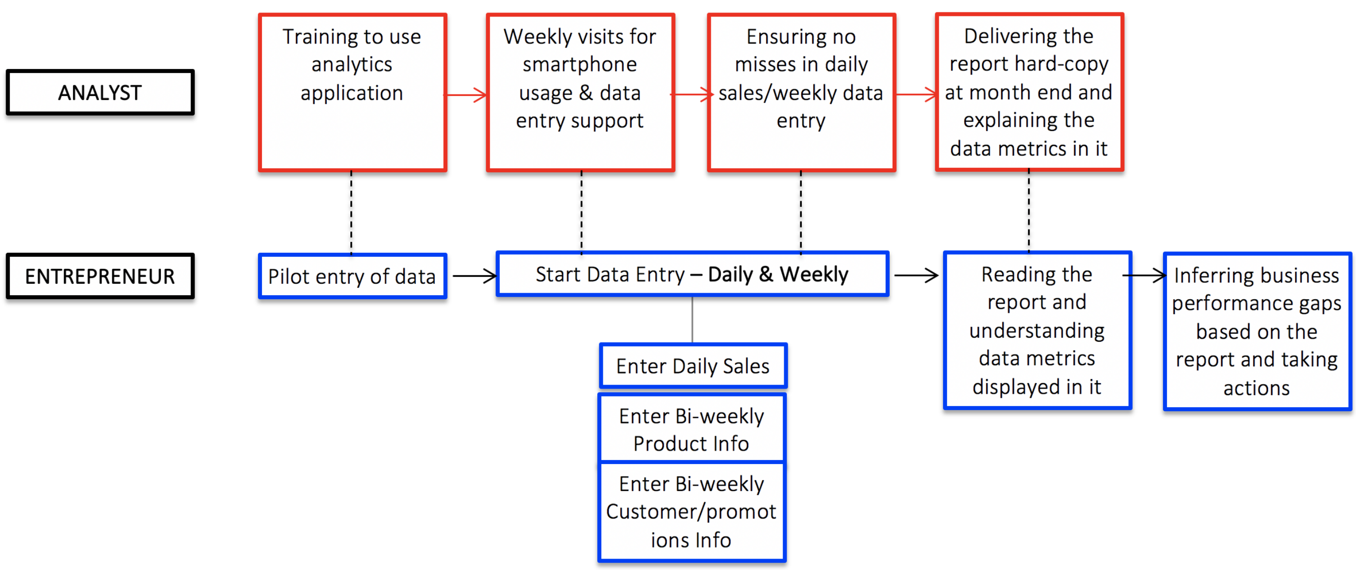


Table A2: Sample Attrition and Survival Checks

	(1) Attrition	(2) Survival
Treatment	0.00400	0.0400*
	(0.0188)	(0.0212)
Placebo	-0.0440***	0.0200
	(0.0130)	(0.0378)
Mean of DV: Control	0.044	0.920
SD of DV: Control	0.205	0.271
Effect Size in SD: Treatment	0.019	0.147
Effect Size in %: Treatment	9.09	4.34
Effect Size in SD: Placebo	-0.214	0.073
Effect Size in %: Placebo	-100	2.17
Obs.	550	550
$\beta\_treat = \beta\_placebo$	0.000	0.577

Notes: This table presents the Attrition and Survival checks for our sample of firms. Column (1) presents the Attition measure (it is 1 if the firm has attrited and 0 if not). Column (2) presents the survival measure which is 1 if firm is operating and 0 if it has shut-down (attrited firms are excluded from the survival checks). Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1$ $p < 0.05^{**}$ $p < 0.01^{***}$

Figure A3: Data Entry Compliance by the entrepreneurs

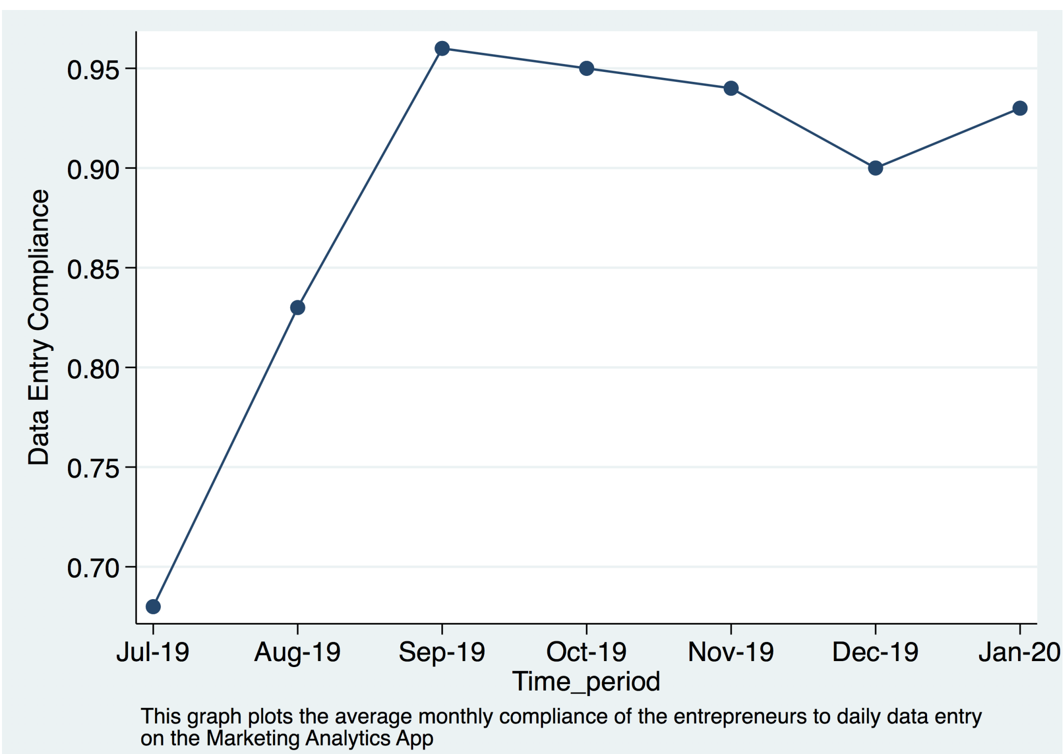


Figure A4: Snapshots from the app depicting the data entry process

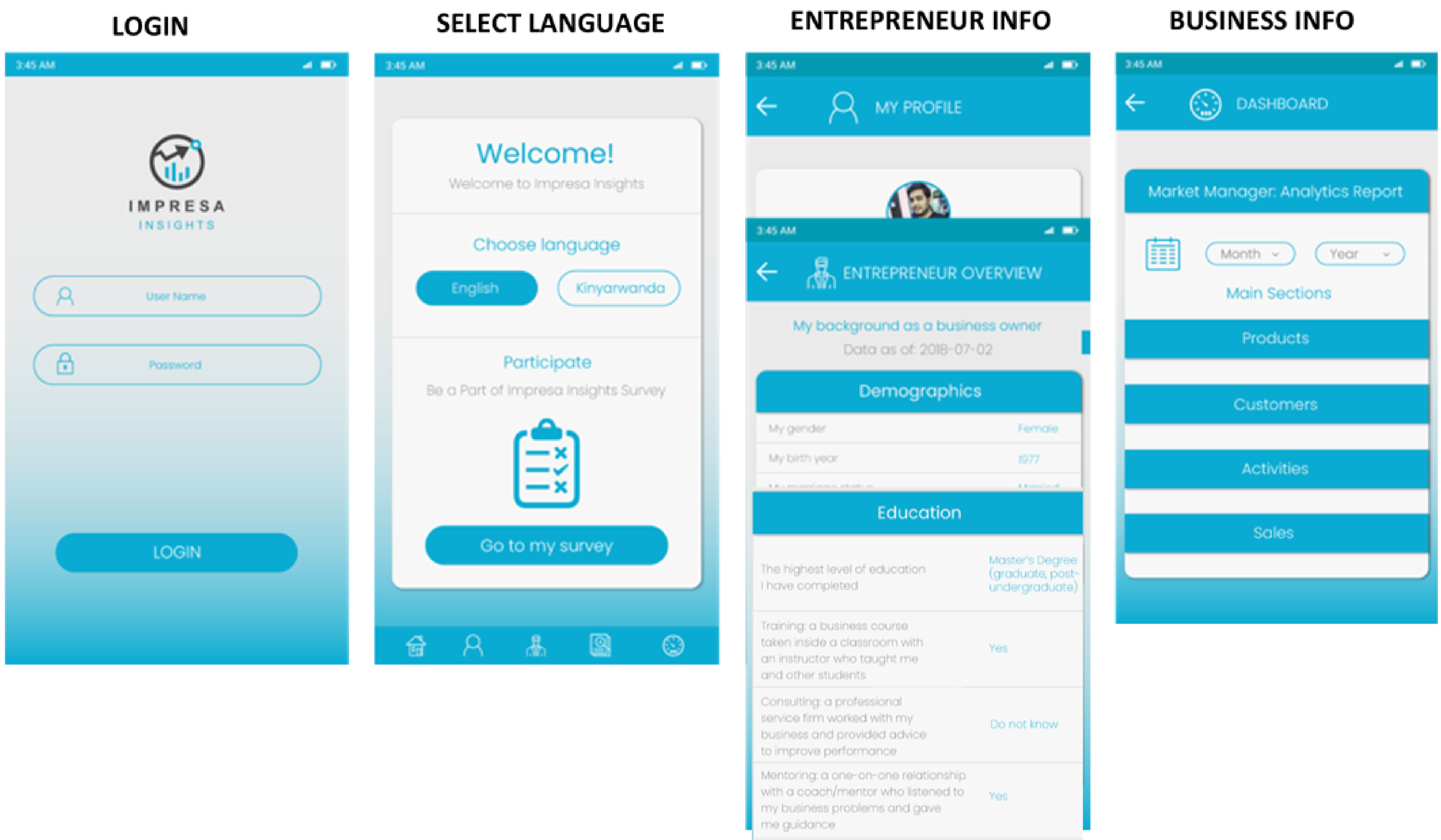


Figure A5: Snapshots from the app depicting the analytics output accessible to the entrepreneur

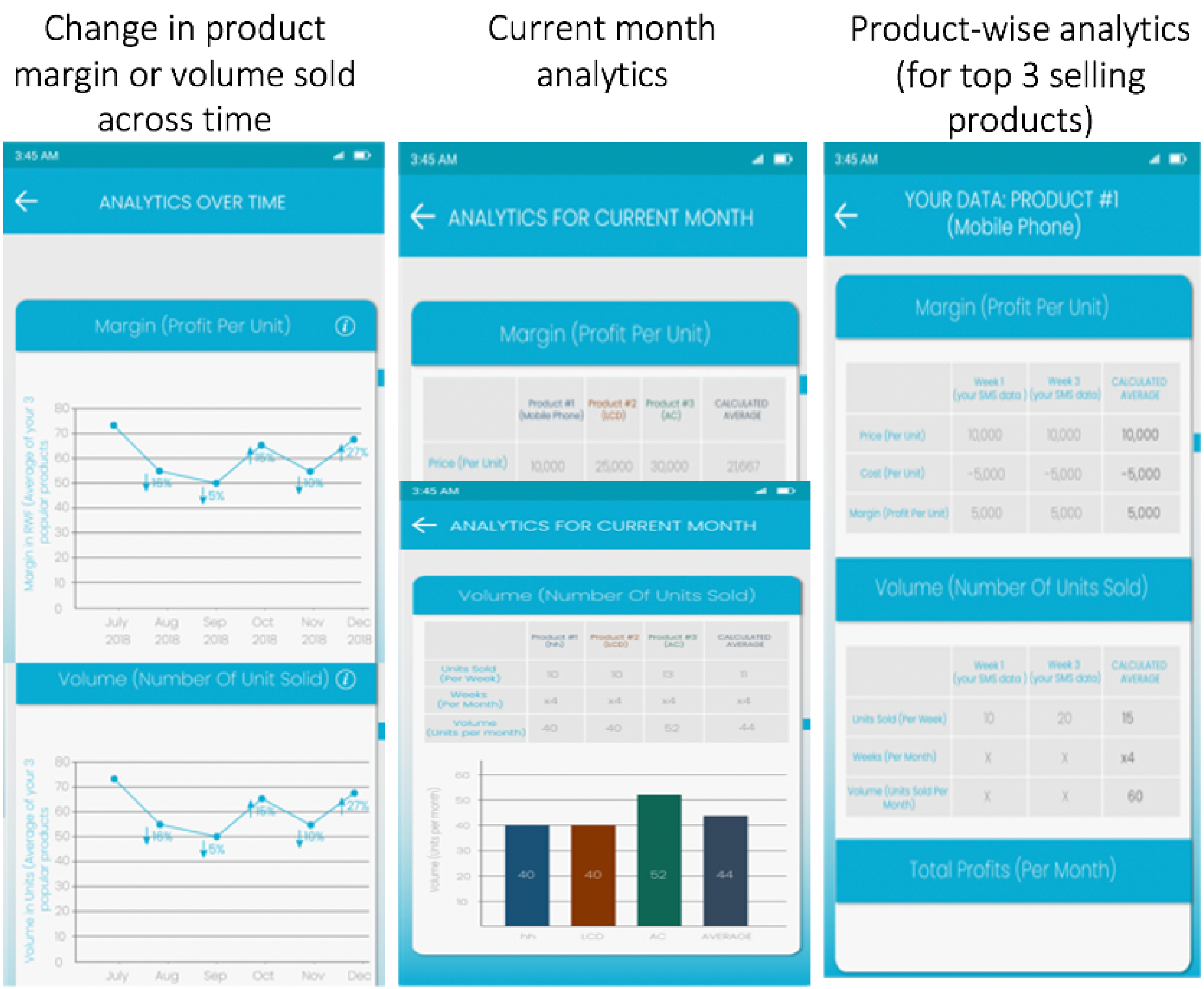
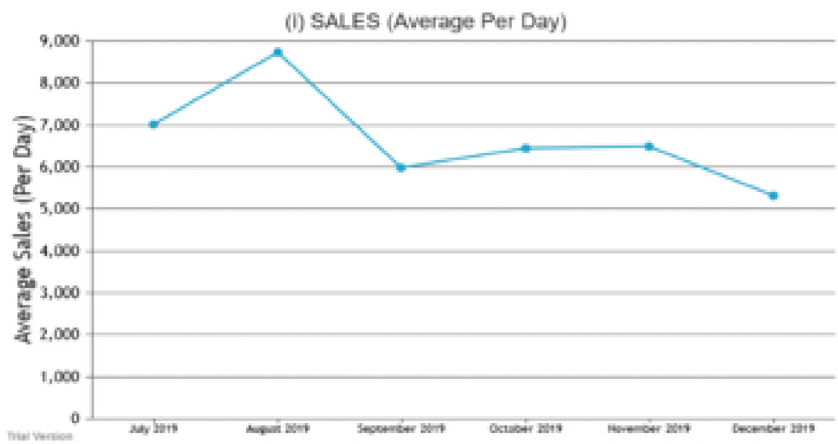


Figure A6: Snapshots from the report depicting the analytics output accessible to the entrepreneur

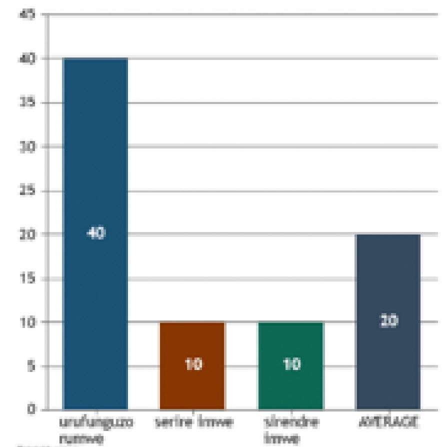
Sales Related Information – Overall and week-wise

MONTHLY SALES (Total Per Month)

Steps to Calculate	Estimated Values
DAYS OPEN (based on question replies)	23
MONTHLY SALES SUB-TOTAL (actual sales entered)	122,300
AVERAGE SALES PER DAY	**5,317**
MISSING RESPONSES	0
IMPUTED DAILY SALES	0
MONTHLY SALES TOTAL	**122,300**



Product, Customer and Activity Based Analytics



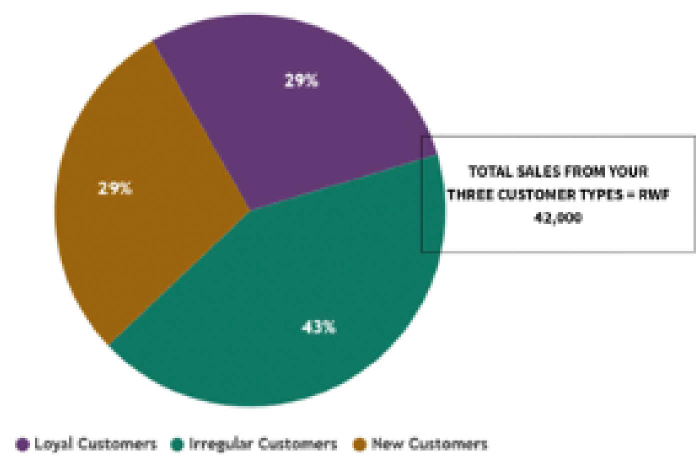


Table A3: Intervention Checks on the use of Analytics Report

	(1) Data Analytics	(2) Customer Analytics	(3) Product Analytics	(4) Sales Analytics
Treatment	2.216***	2.575***	2.436***	2.412***
	(0.107)	(0.101)	(0.109)	(0.0847)
Placebo	0.316	0.396**	0.413*	0.247
	(0.218)	(0.199)	(0.223)	(0.150)
Mean of DV: Control	3.420	2.852	3.403	2.797
SD of DV: Control	1.290	1.069	1.349	0.867
Effect Size in SD: Treatment	1.718	2.409	1.806	2.782
Effect Size in %: Treatment	64.80	90.29	71.60	86.21
Effect Size in SD: Placebo	0.245	0.371	0.307	0.284
Effect Size in %: Placebo	9.237	13.89	12.15	8.812
Obs.	527	527	527	527
$\beta\_treat = \beta\_placebo$	6.01e-18	7.34e-25	1.74e-19	9.94e-39

Notes: This table summarizes the manipulation checks for our marketing analytics intervention on the firms. The analytics report that was presented to the entrepreneurs on a monthly level included each of these sections on customers, products and sales of the firm, hence these results aim to check if the treatment firms have indeed paid attention to these report, as we had intended as a part of our marketing analytics intervention. The results in each column represent the intention-to-treat effects (ITT) of the randomly assigned intervention to a treatment group entrepreneur. Column (1) presents a measure of whether the firms have started paying attention to overall data analytics. Column (2) presents a measure of whether the firms have started paying attention specifically to customer analytics. Column (3) presents a measure of whether the firms have started paying attention specifically to product analytics. Column (4) presents a measure of whether the firms have started paying attention specifically to sales analytics. All regressions include firms that failed as at endline (non-operational with zero monthly sales or profits). Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1$ $p < 0.05^{**}$ $p < 0.01^{***}$

Table A4: Difference in Interview Duration across Experimental Groups

	(1) Interview Duration Analysis
Treatment	0.118
	(0.178)
Placebo	0.413
	(0.346)
Mean of DV: Control	2.512
SD of DV: Control	1.773
Effect Size in SD: Treatment	0.0665
Effect Size in %: Treatment	4.695
Effect Size in SD: Placebo	0.233
Effect Size in %: Placebo	16.46
Obs.	527
$\beta\_treat = \beta\_placebo$	0.404

Notes: This table summarizes analysis for the interview duration for the Endline Survey. Column (1) presents the winsorized measure for time taken from the start to the end of the Endline survey : estimates after winsorizing each 1% on both tails. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1$ $p < 0.05^{**}$ $p < 0.01^{***}$

Figure A7: Model free evidence of main effects - monthly sales

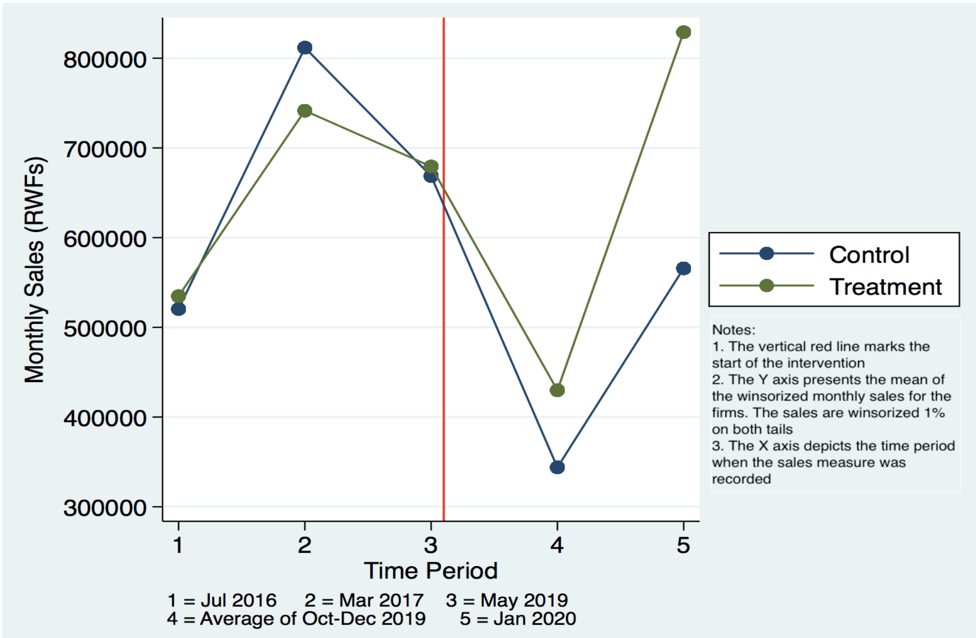
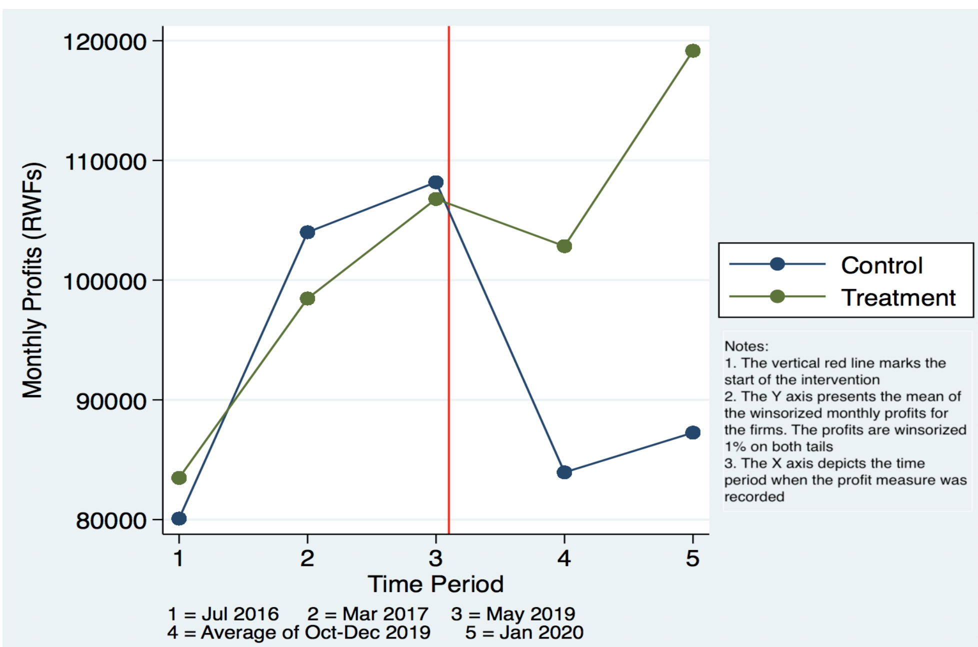


Figure A8: Model free evidence of main effects - monthly profits



Web Appendix B: Robustness to Alternative Mechanisms

Table B1: Balance Checks with the sample at Endline

Variable	(1) **Control (N=239)** Mean [SE]	(2) **Placebo (N=50)** Mean [SE]	(3) **Treatment (N=238)** Mean [SE]	(4) **Total (N=527)** Mean [SE]	(5) **F-test** Joint orthogonality
Capital with which Business was started (RWFs)	3.81e+05	3.27e+05	4.75e+05	4.19e+05	1.480
	[37919.976]	[47560.813]	[56545.972]	[31250.368]	
Paid full-time employees working at the business	0.590	0.340	0.437	0.497	0.838
	[0.134]	[0.109]	[0.067]	[0.069]	
Number of Business Partners	0.092	0.260	0.071	0.099	6.796***
	[0.022]	[0.075]	[0.017]	[0.015]	
Total value of Assets of the Business (RWFs)	2.78e+06	1.68e+06	2.50e+06	2.55e+06	0.750
	[3.87e+05]	[4.20e+05]	[3.95e+05]	[2.53e+05]	
Number of different products sold at the shop	3.464	3.020	3.634	3.499	0.854
	[0.195]	[0.385]	[0.207]	[0.134]	
Number of loyal customers	31.494	24.160	19.496	25.380	1.022
	[8.650]	[4.400]	[1.461]	[4.003]	
Number of new customers	48.230	29.300	28.487	37.518	1.283
	[13.241]	[5.056]	[1.984]	[6.098]	
Basket-size of a loyal customer per visit (RWFs)	3927.568	3602.447	5454.054	4586.629	1.425
	[541.035]	[855.380]	[868.076]	[470.340]	
Basket-size of a new customer per visit (RWFs)	3229.915	6567.708	3108.581	3483.439	1.676
	[717.556]	[4149.033]	[436.671]	[539.432]	
Last month sales (30 days, RWFs)	7.04e+05	6.68e+05	7.29e+05	7.12e+05	0.067
	[65675.656]	[1.65e+05]	[85010.302]	[50959.931]	
Last month profits (30 days, RWFs)	1.27e+05	1.27e+05	1.29e+05	1.28e+05	0.011
	[7162.005]	[16961.824]	[10559.747]	[5979.965]	
Rely on ?gut-feeling? instead of ?data analytics?	2.895	3.040	2.983	2.949	0.428
	[0.080]	[0.187]	[0.083]	[0.055]	
Organizing customers into segments	4.004	3.840	4.050	4.009	1.300
	[0.056]	[0.108]	[0.054]	[0.037]	
Calculating profit margins	4.372	4.440	4.492	4.433	1.971
	[0.046]	[0.091]	[0.039]	[0.029]	
Age of entrepreneur	36.915	35.304	37.609	37.076	1.429
	[0.538]	[0.999]	[0.644]	[0.392]	

*The value displayed for F-tests are the F-statistics. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.*

Table B2: Impact of Marketing Analytics on Firm Sales and Profits for firms surviving till Endline

	DV: Sales				DV: Profits				Overall
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Last Month	IHS Last Month	Typical	IHS Typical	Last Month	IHS Last Month	Typical	IHS Typical	Performance Index
Treatment	268.5**	0.326***	78.07***	0.211***	30.34***	0.231***	17.00**	0.151**	0.262***
	(106.7)	(0.0795)	(26.49)	(0.0696)	(9.124)	(0.0728)	(7.520)	(0.0731)	(0.0730)
Placebo	-131.5	0.0692	-7.013	0.0598	15.04	0.113	4.774	-0.0390	0.0555
	(164.7)	(0.122)	(49.88)	(0.115)	(13.44)	(0.119)	(11.96)	(0.133)	(0.122)
Mean of DV: Control	629.6	13.35	374.0	13.17	95.23	11.84	91.04	11.80	0
SD of DV: Control	1329.6	1.037	305.2	0.886	83.33	0.800	75.98	0.814	0.907
Effect Size in SD: Treatment	0.202	0.314	0.256	0.238	0.364	0.288	0.224	0.185	0.289
Effect Size in %: Treatment	42.65	32.6	20.88	21.1	31.86	23.1	18.68	15.1	
Effect Size in SD: Placebo	-0.0989	0.0667	-0.0230	0.0675	0.180	0.141	0.0628	-0.0479	0.0612
Effect Size in %: Placebo	-20.89	6.92	-1.875	5.98	15.79	11.3	5.244	-3.9	
Obs.	494	494	494	494	494	494	494	494	494
$\beta\_treat = \beta\_placebo$	0.0160	0.0389	0.0923	0.191	0.290	0.328	0.324	0.159	0.0965

Notes: This table summarizes analysis for the main effect of marketing analytics on firm performance outcomes (from baseline to endline). The results in each column represent the effects of the randomly assigned intervention to a treatment group entrepreneur conditional on the firms being operation at the time the Endline Survey was run. Column (1) presents a winsorized measure for Monthly Sales (it refers to the sales for the last full calendar month for the firm, which for our sample was January 2020): estimates after winsorizing each 1% on both tails. Column (2) presents an IHS transformed measure for Monthly Sales: estimates after transforming each with the inverse hyperbolic sine function. Column (3) presents a winsorized measure for Typical Sales (average for the months Nov 2019 to Jan 2020): estimates after winsorizing each 1% on both tails. Column (4) presents an IHS transformed measure for Typical Sales: estimates after transforming each with the inverse hyperbolic sine function. Column (5) presents a winsorized measure for Monthly Profits (it refers to the profits for the last full calendar month for the firm, which for our sample was January 2020): winsorized 1% on both tails. Column (6) presents an IHS transformed measure for Monthly Profits: total monthly profits transformed with the inverse hyperbolic sine function. Column (7) presents a winsorized measure for Typical Profits (average for the months Nov 2019 to Jan 2020): estimates after winsorizing each 1% on both tails. Column (8) presents an IHS transformed measure for Typical Profits: estimates transformed with the inverse hyperbolic sine function. Column (9) presents an index measure for Firm Performance (IHS sales monthly; IHS sales yearly; IHS profits monthly; IHS profits yearly): each of the four measures were standardized and then the average of these values was computed to construct the overall 'performance' index. All regressions exclude firms that failed as at endline (non-operational with zero monthly sales or profits). Any firm growth values in levels (sales; profits) are listed as Rwandan Francs (RWF) in 1000s. The regressions also include: the baseline value of the dependent variable. The IHS results are robust to the scaling of the profits/sales numbers. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1$ $p < 0.05^{**}$ $p < 0.01^{***}$

Table B3: Impact of Marketing Analytics on Firm Sales and Profits - LASSO Regressions

	DV: Sales				DV: Profits				Overall
	(1) Last Month	(2) IHS Last Month	(3) Typical	(4) IHS Typical	(5) Last Month	(6) IHS Last Month	(7) Typical	(8) IHS Typical	(9) Performance Index
Treatment	277.3***	0.821***	87.86***	0.718**	34.83***	0.647**	21.27***	0.566**	0.197**
	(95.99)	(0.310)	(25.54)	(0.285)	(8.146)	(0.277)	(6.788)	(0.257)	(0.0781)
Placebo	-38.71	0.379	13.75	0.267	17.05	0.377	9.003	0.404	0.103
	(162.9)	(0.525)	(43.35)	(0.483)	(13.85)	(0.468)	(11.54)	(0.435)	(0.132)
Mean of DV: Control	565.8	12.23	343.9	12.17	87.26	10.85	83.94	10.91	-8.02e-09
SD of DV: Control	1203.5	3.833	308.9	3.584	84.02	3.374	76.66	3.216	0.982
Effect Size in SD: Treatment	0.230	0.214	0.284	0.200	0.415	0.192	0.277	0.176	0.200
Effect Size in %: Treatment	49.02	82.1	25.55	71.8	39.91	64.7	25.34	56.6	
Effect Size in SD: Placebo	-0.0322	0.0990	0.0445	0.0746	0.203	0.112	0.117	0.126	0.105
Effect Size in %: Placebo	-6.843	37.9	3.997	26.7	19.54	37.7	10.73	40.4	
Obs.	527	527	527	527	527	527	527	527	527
$\beta\_treat = \beta\_placebo$	0.0524	0.400	0.0875	0.351	0.199	0.564	0.288	0.709	0.479

Notes: This table summarizes analysis for the main effect of marketing analytics on firm performance outcomes (from baseline to endline). The results in each column represent the intention-to-treat effects (ITT) of the randomly assigned intervention to a treatment group entrepreneur. Column (1) presents a winsorized measure for Monthly Sales (it refers to the sales for the last full calendar month for the firm, which for our sample was January 2020): estimates after winsorizing each 1% on both tails. Column (2) presents an IHS transformed measure for Monthly Sales: estimates after transforming each with the inverse hyperbolic sine function. Column (3) presents a winsorized measure for Typical Sales (average for the months Nov 2019 to Jan 2020): estimates after winsorizing each 1% on both tails. Column (4) presents an IHS transformed measure for Typical Sales: estimates after transforming each with the inverse hyperbolic sine function. Column (5) presents a winsorized measure for Monthly Profits (it refers to the profits for the last full calendar month for the firm, which for our sample was January 2020): winsorized 1% on both tails. Column (6) presents an IHS transformed measure for Monthly Profits: total monthly profits transformed with the inverse hyperbolic sine function. Column (7) presents a winsorized measure for Typical Profits (average for the months Nov 2019 to Jan 2020): estimates after winsorizing each 1% on both tails. Column (8) presents an IHS transformed measure for Typical Profits: estimates transformed with the inverse hyperbolic sine function. Column (9) presents an index measure for Firm Performance (IHS sales monthly; IHS sales yearly; IHS profits monthly; IHS profits yearly): each of the four measures were standardized and then the average of these values was computed to construct the overall 'performance' index. All regressions include firms that failed as at endline (non-operational with zero monthly sales or profits). Any firm growth values in levels (sales; profits) are listed as Rwandan Francs (RWF) in 1000s. The LASSO regressions include the baseline value of the dependent variable and: twenty controls for entrepreneur characteristics at baseline (these include - employee mindset, use of data Vs intuition, education, prior business ownership, prior training/consulting/mentorship in business, entrepreneur's age); and nineteen controls for business characteristics at baseline (provision of goods Vs services, start-up capital, number of employees, asset value, number of products offered, margin of products, number of customers - new vs loyal, basket size of customers - new vs loyal, firm's sales and profits at baseline). The IHS results are robust to the scaling of the profits/sales numbers. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1$ $p < 0.05^{**}$ $p < 0.01^{***}$

Table B4: Robustness Checks for Main Effects - Including Controls

	DV: Sales				DV: Profits				Overall
	(1) Last Month	(2) IHS Last Month	(3) Typical	(4) IHS Typical	(5) Last Month	(6) IHS Last Month	(7) Typical	(8) IHS Typical	(9) Performance Index
Treatment	261.7**	0.841**	92.38***	0.756**	34.25***	0.673**	20.83***	0.583**	0.203**
	(105.4)	(0.346)	(25.98)	(0.320)	(8.076)	(0.304)	(6.886)	(0.287)	(0.0878)
Placebo	-26.69	0.258	14.79	0.114	17.86	0.263	7.458	0.248	0.0640
	(123.2)	(0.579)	(39.39)	(0.558)	(11.86)	(0.522)	(10.60)	(0.455)	(0.146)
Mean of DV: Control	565.8	12.23	343.9	12.17	87.26	10.85	83.94	10.91	-8.02e-09
SD of DV: Control	1203.5	3.833	308.9	3.584	84.02	3.374	76.66	3.216	0.982
Effect Size in SD: Treatment	0.217	0.219	0.299	0.211	0.408	0.200	0.272	0.181	0.207
Effect Size in %: Treatment	46.25	84.1	26.86	75.6	39.25	67.3	24.81	58.3	
Effect Size in SD: Placebo	-0.0222	0.0673	0.0479	0.0318	0.213	0.0778	0.0973	0.0770	0.0651
Effect Size in %: Placebo	-4.717	25.8	4.300	11.4	20.47	26.3	8.884	24.8	
Obs.	527	527	527	527	527	527	527	527	527
$\beta\_treat = \beta\_placebo$	0.0182	0.318	0.0494	0.251	0.189	0.439	0.220	0.457	0.340

Notes: This table summarizes analysis for the main effect of marketing analytics on firm performance outcomes (from baseline to endline). The results in each column represent the intention-to-treat effects (ITT) of the randomly assigned intervention to a treatment group entrepreneur, while including covariates on entrepreneur and business characteristics. Column (1) presents a winsorized measure for Monthly Sales (it refers to the sales for the last full calendar month for the firm, which for our sample was January 2020): estimates after winsorizing each 1% on both tails. Column (2) presents an IHS transformed measure for Monthly Sales: estimates after transforming each with the inverse hyperbolic sine function. Column (3) presents a winsorized measure for Typical Sales (average for the months Nov 2019 to Jan 2020): estimates after winsorizing each 1% on both tails. Column (4) presents an IHS transformed measure for Typical Sales: estimates after transforming each with the inverse hyperbolic sine function. Column (5) presents a winsorized measure for Monthly Profits (it refers to the profits for the last full calendar month for the firm, which for our sample was January 2020): winsorized 1% on both tails. Column (6) presents an IHS transformed measure for Monthly Profits: total monthly profits transformed with the inverse hyperbolic sine function. Column (7) presents a winsorized measure for Typical Profits (average for the months Nov 2019 to Jan 2020): estimates after winsorizing each 1% on both tails. Column (8) presents an IHS transformed measure for Typical Profits: estimates transformed with the inverse hyperbolic sine function. Column (9) presents an index measure for Firm Performance (IHS sales monthly; IHS sales yearly; IHS profits monthly; IHS profits yearly): each of the four measures were standardized and then the average of these values was computed to construct the overall 'performance' index. All regressions include firms that failed as at endline (non-operational with zero monthly sales or profits). Any firm growth values in levels (sales; profits) are listed as Rwandan Francs (RWF) in 1000s. The regressions include: the baseline value of the dependent variable. In addition, we also include twenty controls for entrepreneur characteristics at baseline (these include - employee mindset, use of data Vs intuition, education, prior business ownership, prior training/consulting/mentorship in business, entrepreneur's age); and nineteen controls for business characteristics at baseline (provision of goods Vs services, start-up capital, number of employees, asset value, number of products offered, margin of products, number of customers - new vs loyal, basket size of customers - new vs loyal, firm's sales and profits at baseline) The IHS results are robust to the scaling of the profits/sales numbers. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1$ $p < 0.05^{**}$ $p < 0.01^{***}$

Table B5: Change in Business Time

	(1) Percent_Days_Open
Month_num	-0.0000184
	(0.00260)
Obs.	1428

Notes: This table summarizes, change in percentage of days for which the treatment entrepreneur's shop was open across-months (July to Dec 2019). We use a panel data of month on month percentage of days for which each of the 238 treatment shops (post attrition) was open to doing business. The regression includes firm level fixed effects and standard errors are clustered at firm level. P-values are highlighted as: $p < 0.1$ p $<0.05^{**}$ $p < 0.01^{***}$

Table B6: Correlation between sales submitted in the App and reported in the Endline Survey

	(1) Endline Survey Sales
App Sales	1.004^{***}
	(0.0538)
Obs.	225
$\beta\_AppSales = 1$	0.935
$Constant = 0$	0.890

Notes: This table summarizes analysis for the comparison between the last month sales reported in the Endline Survey and the Analytics Application. The treatment entrepreneurs entered daily sales in the analytics application which was added for the last month (Jan 2020) in order to calculate the sales for the full month (as per the application). The analysis is conducted for 225 of the 238 treatment entrepreneurs, as the remaining did not enter data in the analytics application for the month of interest. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1$ p $<0.05^{**}$ $p < 0.01^{***}$

Table B7: Robustness Checks for Main Effects - Additional Evidence

	(1) Product Sales Metric	(2) IHS Product Sales Metric	(3) Product Profits Metric	(4) IHS Product Profits Metric	(5) Customer Purchase	(6) IHS Customer Purchase	(7) Referrals	(8) IHS Referrals
Treatment	52.55	0.546*	3.617	0.499*	112.0**	0.775**	0.268	0.203**
	(37.99)	(0.286)	(5.630)	(0.255)	(55.47)	(0.304)	(0.871)	(0.101)
Placebo	-5.374	0.242	1.217	0.381	118.3	0.495	0.0702	0.0592
	(35.83)	(0.488)	(7.410)	(0.435)	(114.1)	(0.524)	(1.741)	(0.173)
Mean of DV: Control	150.3	11.05	32.64	9.627	277.5	11.47	5.548	1.709
SD of DV: Control	349.6	3.492	58.29	3.120	494.8	3.690	10.40	1.151
Effect Size in SD: Treatment	0.150	0.156	0.0620	0.160	0.226	0.210	0.0258	0.176
Effect Size in %: Treatment	34.98	4.947	11.08	5.187	40.35	6.756	4.829	11.85
Effect Size in SD: Placebo	-0.0154	0.0694	0.0209	0.122	0.239	0.134	0.00675	0.0514
Effect Size in %: Placebo	-3.577	2.193	3.728	3.954	42.63	4.317	1.265	3.464
Obs.	527	527	527	527	527	527	527	527
$\beta\_treat = \beta\_placebo$	0.161	0.515	0.753	0.775	0.958	0.579	0.907	0.402

Notes: This table summarizes additional evidence for the main effect of marketing analytics on firm performance outcomes (from baseline to endline) by analyzing the product and customer related outcome variables. The results in each column represent the intention-to-treat effects (ITT) of the randomly assigned intervention to a treatment group entrepreneur. Column (1) presents a winsorized measure for Product Sales (it refers to the sum-product of the prices and volumes of the top 3 selling products of the business for the last full week for the firm): estimates after winsorizing each 1% on both tails. Column (2) presents an IHS transformed measure for Product Sales: estimates after transforming each with the inverse hyperbolic sine function. Column (3) presents a winsorized measure for Product Profit (it refers to the sum-product of the margins and volumes of the top 3 selling products of the business for the last full week for the firm): estimates after winsorizing each 1% on both tails. Column (4) presents an IHS transformed measure for Product Profit: estimates after transforming each with the inverse hyperbolic sine function. Column (5) presents a winsorized measure for Customer Purchase (it refers to the product of the total number of customers and their average basket size per visit, for the last full calendar month for the firm which for our sample was January 2020): winsorized 1% on both tails. Column (6) presents an IHS transformed measure for Customer Purchase: estimates transformed with the inverse hyperbolic sine function. Column (7) presents a winsorized measure for Customer Referrals (average for the number of customers who have recommended the firm or its products to potential customers): estimates after winsorizing each 1% on both tails. Column (8) presents an IHS transformed measure for Customer Referrals: estimates transformed with the inverse hyperbolic sine function. All regressions include firms that failed as at endline (non-operational with zero monthly sales or profits). Any firm growth values in levels (sales; profits) are listed as Rwandan Francs (RWF) in 1000s. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1$ $p < 0.05^{**}$ $p < 0.01^{***}$

Table B8: Robustness Checks for Main Effects - Attrition Bounding

	DV: Sales				DV: Profits			
	(1) Sales Growth Bound 1	(2) Sales Growth Bound 2	(3) Sales Growth Bound 3	(4) Sales Growth Bound 4	(5) Profits Growth Bound 1	(6) Profits Growth Bound 2	(7) Profits Growth Bound 3	(8) Profits Growth Bound 4
Treatment	228.4**	247.5**	261.9***	249.6**	28.96***	31.98***	37.25***	35.59***
	(100.1)	(96.26)	(96.73)	(96.93)	(8.898)	(9.019)	(10.21)	(10.27)
Placebo	-57.65	-82.54	-81.81	-93.83	20.73	17.44	18.38	16.78
	(131.9)	(143.6)	(147.2)	(147.9)	(13.51)	(12.85)	(12.96)	(13.00)
Mean of DV: Control	540.9	568.4	568.4	580.5	83.42	88.23	88.23	89.86
SD of DV: Control	1182.4	1185.7	1185.7	1197.2	84.08	85.31	85.31	88.37
Effect Size in SD: Treatment	0.193	0.209	0.221	0.209	0.344	0.375	0.437	0.403
Effect Size in %: Treatment	42.23	43.54	46.08	43.00	34.71	36.24	42.23	39.61
Effect Size in SD: Placebo	-0.0488	-0.0696	-0.0690	-0.0784	0.247	0.204	0.215	0.190
Effect Size in %: Placebo	-10.66	-14.52	-14.39	-16.16	24.84	19.76	20.84	18.68
Obs.	550	550	550	550	550	550	550	550
$\beta\_treat = \beta\_placebo$	0.0426	0.0293	0.0269	0.0276	0.569	0.294	0.185	0.187

Notes: This table summarizes robustness analysis for the main effect of marketing analytics on firm performance outcomes (from baseline to endline). We show treatment effects under different assumptions on attrition. Column (1) assigns all attriters a sales growth of zero. Column (2) assigns all attriters the average sales growth of the control group. Column (3) assigns all control attiters the sales growth rate of the control group and all treatment attriters a sales growth of zero. Column (4) assigns all control attriters the average sales growth of the treatment groups and all treatment attriters a sales growth of zero. Column (5) assigns all attriters a profit growth of zero. Column (6) assigns all attriters the average profit growth of the control group. Column (7) assigns all control attiters the profit growth rate of the control group and all treatment attriters a profit growth of zero. Column (8) assigns all control attriters the average profit growth of the treatment groups and all treatment attriters a profit growth of zero. All regressions include firms that failed as at endline (non-operational with zero monthly sales or profits). Any firm growth values in levels (sales; profits) are listed as Rwandan Francs (RWF) in 1000s. The regressions also include: the baseline value of the dependent variable. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1$ $p < 0.05^{**}$ $p < 0.01^{***}$

Table B9: Robustness Checks for Main Effects - Average Treatment Effect on the Treated

	DV: Sales		DV: Profits		Overall
	(1)	(2)	(3)	(4)	(5)
	Last Month	Typical	Last Month	Typical	Performance Index
Treatment	353.8***	112.1***	40.74***	25.22***	0.354***
	(105.4)	(27.35)	(9.465)	(7.642)	(0.0651)
Mean of DV: Control	565.8	343.9	87.26	83.94	0
SD of DV: Control	1203.5	308.9	84.02	76.66	0.982
Effect Size in SD: Treatment	0.294	0.363	0.485	0.329	0.360
Effect Size in %: Treatment	62.54	32.60	46.69	30.04	
Obs.	439	439	439	439	439

Notes: This table summarizes analysis for the main effect of marketing analytics on firm performance outcomes (from baseline to endline). The results in each column represent the average treatment-on-treated effects (ATT) of the randomly assigned intervention to a treatment group entrepreneur, computed via 2SLS, with treatment adoption (defined as signing-up on the analytics app) and compliance (defined as entering data for at least one week for each intervention month between July - Dec 2019) instrumented by the random treatment assignment. Column (1) presents a winsorized measure for Monthly Sales (it refers to the sales for the last full calendar month for the firm, which for our sample was January 2020): estimates after winsorizing each 1% on both tails. Column (2) presents a winsorized measure for Typical Sales (average for the months Nov 2019 to Jan 2020): estimates after winsorizing each 1% on both tails. Column (3) presents a winsorized measure for Monthly Profits (it refers to the profits for the last full calendar month for the firm, which for our sample was January 2020): winsorized 1% on both tails Column (4) presents a winsorized measure for Typical Profits (average for the months Nov 2019 to Jan 2020): estimates after winsorizing each 1% on both tails. Column (5) presents an index measure for Firm Performance (sales monthly; sales yearly; profits monthly; profits yearly): the IHS transforms of each of the four measures were standardized and then the average of these values was computed to construct the overall 'performance' index. All regressions include firms that failed as at endline (non-operational with zero monthly sales or profits). Any firm growth values in levels (sales; profits) are listed as Rwandan Francs (RWF) in 1000s. The regressions also include: the baseline value of the dependent variable. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1$ $p < 0.05^{**}$ $p < 0.01^{***}$

Table B10: Robustness Checks for Main Effects - Impact on Control Firms if There is Competing Treatment Firm Nearby

	DV: Monthly Sales		DV: Monthly Profits		Overall	
	(1)	(2)	(3)	(4)	(5)	(6)
	Average	IHS	Average	IHS	Performance Index	Average Performance Index
Close Treatment Competitor	84.78	0.273	13.10	0.227	0.0913	0.0682
	(65.61)	(0.615)	(10.18)	(0.578)	(0.174)	(0.174)
Mean of DV: Control	330.6	11.95	70.76	10.66	-0.104	-0.0887
SD of DV: Control	303.1	3.880	56.95	3.442	1.065	1.066
Effect Size in SD: Competitor	0.280	0.0704	0.230	0.0659	0.0858	0.0640
Effect Size in %: Competitor	25.64	2.287	18.51	2.128	-87.44	-76.98
Obs.	239	239	239	239	239	239

Notes: This table summarizes analysis for the spillover effect on the control firms of having a treatment firm nearby. The regression is run on the sample of control firms, and provide effect on the control firm of having at least one treatment firm nearby (within 1 km of Haversine distance from the focal firm) which belongs to the same sector (pre coded industry variable) as the focal firm. Column (1) presents a winsorized measure for Average Monthly Sales (it refers to the average sales for the last full calendar month for the firm, which for our sample was January 2020) and the sales for a typical month (average for the months Oct 2019 to Dec 2020) : estimates after winsorizing each 1% on both tails. Column (2) presents an IHS transformed measure for Average Monthly Sales: estimates after transforming the average monthly sales with the inverse hyperbolic sine function. Column (3) presents a winsorized measure for Average Monthly Profits (it refers to the average profits for the last full calendar month for the firm, which for our sample was January 2020) and the profits for a typical month (average for the months Oct 2019 to Dec 2020) : estimates after winsorizing each 1% on both tails. Column (4) presents an IHS transformed measure for Average Monthly Profits: estimates after transforming the average monthly profits using the inverse hyperbolic sine function. Column (5) presents an index measure for Firm Performance (IHS last month's sales; IHS typical sales monthly; IHS last month's profits; IHS typical profits monthly): each of the four measures were standardized and then the average of these values was computed to construct the overall 'performance' index. Column (6) presents an index measure for Firm Performance (using IHS average sales monthly; IHS average profits monthly): each of the two measures were standardized and then the average of these values was computed to construct the overall 'performance' index. All regressions include firms that failed as at endline (non-operational with zero monthly sales or profits). Any firm growth values in levels (sales; profits) are listed as Rwandan Francs (RWF) in 1000s. The regressions also include: the baseline value of the dependent variable. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1$ $p < 0.05^{**}$ $p < 0.01^{***}$

Table B11: Impact of Marketing Analytics on Other Activities of the Business

	(1) Machine Maintenance	(2) Process Efficiency	(3) Competition Research	(4) Receipts or Invoices	(5) Other Activities Index
Treatment	-0.00840	0.0221	0.0189	0.000141	0.00850
Placebo	0.0659	0.0722	0.0463	-0.0135	0.0430
Mean of DV: Control	0.294	0.268	0.494	0.0335	0.272
SD of DV: Control	0.457	0.444	0.501	0.180	0.247
Effect Size in SD: Treatment	-0.0184	0.0499	0.0377	0.000780	0.0343
Effect Size in %: Treatment	-2.857	8.265	3.824	0.420	3.124
Effect Size in SD: Placebo	0.144	0.163	0.0924	-0.0747	0.174
Effect Size in %: Placebo	22.40	26.97	9.373	-40.25	15.82
$\beta\_treat = \beta\_placebo$	0.317	0.495	0.725	0.555	0.398

Notes: This table summarizes analysis for the effect of marketing analytics on activities that the business performs (from baseline to endline). The results in each column represent the intention-to-treat effects (ITT) of the randomly assigned intervention to a treatment group entrepreneur. These results are based on a Yes/No response on each of these activities, as recorded by the auditors in-person review of each of the firms in our sample. Column (1) presents a measure for conducting preventative maintenance on the machines, equipments or tools used by the business. Column (2) presents a measure for replacing any processes, procedures, techniques or methods used in the business with new approaches that are more efficient. Column (3) presents a measure for conducting market research on competition and suppliers. Column (4) presents a measure of provision of receipts or invoices to customers once they have made a purchase. Column (5) presents an index for these 'other' activities performed by the business. This overall index is calculated by taking an average of each of the four business measures. All regressions include firms that failed as at endline (non-operational with zero monthly sales or profits, and their business activity response is considered to be 'no'. All regression exclude the firms that attrited during the Endline survey round. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1$ $p < 0.05^{**}$ $p < 0.01^{***}$

Table B12: Mediation Analysis - Analytical Ability of the Entrepreneur

	Estimate	p-value
ACME	0.1578	2e-16***
ADE	0.0229	0.814
Total Effect	0.1807	0.008***
Proportion Mediated	0.8735	0.008***

Non parametric bootstrap confidence intervals with the Percentile Method
ACME = Average Causal Mediation Effect and ADE = Average Direct Effect
Sample size = 527, simulations = 1000
P-values are highlighted as: $p < 0.1$ $p < 0.05^{**}$ $p < 0.01^{***}$

Table B13: Mediation Analysis - Analytical Activity of the Entrepreneur

	Estimate	p-value
ACME	0.6597	2e-16***
ADE	-0.4709	2e-16***
Total Effect	0.1889	0.008***
Proportion Mediated	3.493	0.008***

Non parametric bootstrap confidence intervals with the Percentile Method
ACME = Average Causal Mediation Effect and ADE = Average Direct Effect
Sample size = 527, simulations = 1000
*P-values are highlighted as: p < 0.1 p <0.05** p < 0.01****

Figure B1: Post COVID-19, Qualitative Survey

Firm	Data Recording	Data Analytics	Role of Data Analytics during COVID-19 related shocks
General Merchandise Store (Retail Trade)	I keep a record of the number of units sold per product and of my competitors' prices. I have it all written in a notebook that I maintain. I also record the profit I make per month, people who owe me money and how much. In addition I keep a separate record of how much I received on mobile money everyday at the end of the day.	At the end of each week, I review the data that I have collected for my business and analyze it to make projections about customer purchase patterns. It also hels me to know when to add new products or increase my stock. I can also see how a product is being bought, and understand if customers are liking it or not, accordingly I can choose to replace it with a new product or a similar product from a different brand. For example, I have noticed that my consumers make a lot of purchases through mobile payments on Friday, just before going for the weekend so I ensure I have adequate stock for Fridays. I also noticed that at the beginning of the school year or when schools re-open after vacations, I get a lot of consumers seeking school-related products so, I stock-up accordingly with things that students need especially boarding-school students.	Yes most definitely data analytics helped me get through the lockdowns, I was able to ensure that I reach my monthly target of sales by doing home delivery, and I was able to pay for the rent because I knew how much to sell to keep operating my business.
Music and Video Entertainment (Services)	Yes, I maintain an excel sheet with the daily sales (number of units) and also the cost of the internet that I buy to be able to download movies or videos that I sell at my store.	Yes I do analyze the data I maintain in the excel sheet. In fact it was by analyzing this data that I realized I was not making much money through movie rentals because I had to invest a lot in downloading them. So now, I have completely shifted my business to music rental only. I also have seen from my sales and product data that people now prefer more Rwandan songs, so I am no longer spending time and data downloading other songs. Further, as a result of COVID lockdowns, young people are not going to school and I noticed an icrease in young customers at my store. I am now catering to their tastes better by creating lists of songs that the youth enjoy.	No it did not help my business during COVID shocks

Firm	Data Recording	Data Analytics	Role of Data Analytics during COVID-19 related shocks
Garment Styling and Alterations (Services)	I record the sales revenue, number of clothes like dresses, shirts etc that I style, and my monthly profit. I have a book in which I put in all my orders and sales on a daily basis. I also record the advance that people pay, the number of orders and their deadlines.	I periodically look through my recorded sales and order data. That way, I have been able to check which of my loyal customers have not been ordering from me, and I call them to ensure that I retain them. Based on the total number of orders that I have, I can purchase more fabric, and also know if I can give some orders to my colleagues and pay them (if I am facing too many urgent orders). I had also started to target customers who had weddings, and that way I was able to make a lot of money. But now due to COVID, since wedding ceremonies are not allowed to be held, that sale has stopped.	No, the lockdown was so unpredictable that it was extremely difficult to pre-empt anything even based on the information I have about my business.
Kitchen Supplies Store (Retail Trade)	I record stock data for each of my products, weekly sales data, prices of the product and the addresses of my loyal customers. Every Saturday I record this data on a fresh sheet of paper. The customer order data is recorded for my - loyal, irregular and new customers, along with product information, on a weekly basis.	I have to analyse the orders that have been made and their volume, to understand how much I need to purchase as raw material stock. According to my sales data, I know what to stock in good quantity. Analyzing the number of units ordered for each product, helps me understand what people like more, and this helps me in ensuring that I always have those products at my shop. For example, customers have started ordering fruits (especially citric fruits) a lot of because of COVID, and I have ensured that every weekend I have enough stock of lemons, lime, oranges etc.	Yes, I had a full list of my loyal customers and the last time that they had purchased from my store, along with the products they generally purchase. So, I continued to focus on my loyal customers by actively calling them, taking orders from them and ensuring timely delivery to their houses.
Charcoal shop (Retail trade)	I keep a dedicated notebook. In which I write down the daily sales revenue of my shop, number of units sold, monthly profits, number of pending orders and loyal customers information. In the same bookI also note down all the expenses that I have (including loading and offloading costs), man power etc.	Analyzing the sales data has proved to be of a great help for me. Due to the decrease in the prices of gas the sale of charcoal has taken a hit, and hence I have suspended some casual workers who used to helped me in the past. But because I knew that my sales are going to reduce, and my new projected margins and sales were not able to cover the costs of the casuals, I had to let them go.	Yes analytics did help me, because I had a record of my loyal customers and could find them in their homes, and deliver to them my product.

Firm	Data Recording	Data Analytics	Role of Data Analytics during COVID-19 related shocks
Hair Styling Salon (Services)	I just record the number of clients that visited my shop per day.	I haven't been using the data lately as there is not much I can do with it in my business.	Yes analytics did help me a lot. Actually I maintained a record of my loyal customers along with their contact information and preferences. So during lockdown, I was able to schedule their appointments over the phone and cater to them through home-visits. That kept my sales a little bit stable.
Shoe Repairing Shop (Services)	I don't do any data recording at all now. I had to give up on the physical space of my shop due to loss of business during the COVID lockdowns. I am sitting in the street now to run my services, and cannot find time for any data work.	I don't work with any business data as it is very difficult to that in my current business situation.	No it did not.
Restaurant (Services)	I have a business book that I have dedicated for recording sales data, stock data, expenses and customers' contact. I have also started monthly subscription service for my customers and record the same along with when to expect money from each of my subscription customer.	I review my data almost on a daily basis. By comparing the prices of my stock with the prices charged by other similar stores in the market, I know how much I can charge my customers. The same information also gives me a picture of how much projected profit I can make. Total sales per unit data tells me which products I can invest in more. It also tells me what types of products different customer segments like, accordingly I customize the food plate for breakfast or for lunch that I offer.	No, I only worked with the hospitals to provide them with tiffins during the lockdown.
Garment Shop (Retail Trade)	I have a book that I use to record all my sales, number of orders received, the advance payment I have received and how much people owe me (credit). I also periodically account for the payments I receive through Mobile money, cash and money in the bank to see my total sales.	I set monthly sales targets for my business and analyzing the recorded data help me to know if I have been able to reach my sales goal or how much was I short. I can accordingly reach out to some of the clients who haven't been showing up and see if they can come to the shop, to ensure I reach my sales goal. This also helps me to know if my sales will be enough to keep the helpers who help me to cut the fabrics, to iron the clothes etc, or if I will have to plan to do these tasks myself for that month. The projections of sales also help me order adequate fabrics.	Yes data analytics helped me as I had a list of my loyal customers, and I have been contacting them and providing home based services.

Firm	Data Recording	Data Analytics	Role of Data Analytics during COVID-19 related shocks
Electronics Repair Shop (Services)	I record daily sales and daily expenses in a notebook. This helps me calculate the business profit on a daily basis. I periodically also. Record some key stock information.	I just analyze my profit on a daily basis. Apart from that, there are some spare parts that I have to always ensure I have at all times. I keep monitoring there stock level information and always avoid to have zero stocks on those items.	Yes the data helped me a lot. As I already knew my monthly sales estimations, in the second week of the lockdown and curfew, I had already understood that I won't be able to pay the rent and the electricity for my shop. So, I decided to leave the shop and operate from my home.
Automobile Repair Shop (Services)	I just maintain a list of products I stock, along with their prices, but we don't change that very often.	I don't really invest much time analyzing the data, because I don't record much of it anyway.	No, I don’t think so.
Garments Shop (Retail Trade)	I record the new stock purchases and their sales. Everyday I keep track of any new customers that come to my store, for them I record what s/he pays and what products they purchased. I do this only for my new clients who come to my shop for the first time. I also record business expenses like rent and electricity. I use an app in my phone to record all this information (it is a different app, not the one which was provided to us as a part of the project).	I check my weekly sales to see if I need to increase the price of my products for the rest of the month in order to get to my monthly profit target. Almost everyday, I also check on how the month has been so far and what the new sale mean is for me. Depending on how the current sales are, I come to know if I will need new stocks before the end of the month and I order accordingly.	Yes, during the pandemic, I have only focused on the most popular products, and that is all that I have sold. This helped me to sell my whole stock quickly and buy new ones. This has also helped me in doing home-deliveries because I knew the contact information as well as product-prefeences of my loyal customers. I called them and they purchased all the stock I had.

Firm	Data Recording	Data Analytics	Role of Data Analytics during COVID-19 related shocks
Grocery Store (Retail Trade)	I capture daily sales data in my business book. This also includes what I have sold, its quantity, type of product, and the cost of purchase to my business. I also record the phone numbers and other information of any new customer that comes to my shop. I do that so that I can call them as a reminder to come back and visit/purchase from my shop.	I can't do anything in my business without anakyzing the data beforehand. I firstly check the status of my business using the recorded data. Everyday, I go through my daily sales and do an audit to see if I didn't lose any money. When I am thinking of giving people credit or when I am thinking of introducing a new product in my store, I go through the same numbers and consider that data.	Yes, a lot, because some of our products can be damaged when they are not purchased directly, or they just get expired. Having in mind the average of my monthly or weekly sales was very helpful to minimize that loss. I also knew my loyal customers and the quantity they generally like to purcahse. So, everytime that I purchased new stock, I considered that.
Dairy Products and Hot Beverages Store (Retail Trade)	I have been keeping a close tab on my daily sales data ever since the project began, I am still continuing to do it as it has helped me a lot. I record sales data for every customer who purchases something at my shop. I also take note of my monthly expenses like rent, raw-materials and employee payments.	I make a projection of sales for each day based on my past sales numbers. This helps me in knowing how much I should purchase as stock, and minimizes over-stocking or under-stocking as I purchase the products according to my data-based estimate of how much the consumption is going to be for the day. During sunny periods people I have noticed that people don't drink hot milk or tea. So I have started to provide juice as well at my shop, and I accordingly reduce the number of casual workers for those days as we don't have to prepare tea.	Yes, during the COVID lockdown, using analytics I had identified my loyal customers and I knew what they buy everytime from my store. After the lockdown, I have invested a lot in those people and asked them what they would like to get on reduced price because most of them didn't have a lot of money. I have managed to retain them until now and they form a bigger percentage of my sales today,