Innovation and Entrepreneurship : Empirical evidence using micro data

Leonard Sabetti

Ecole Doctorale des Sciences Economiques, Juridiques, Politiques et de gestion
Centre d'Etudes et de Recherche sur le Développement International (CERDI)

Université Clermont Auvergne, CNRS, IRD, CERDI, F-63000 Clermont-Ferrand, France

L'innovation et l'esprit d'entreprise

Preuves empiriques à l'aide de microdonnées

Thèse présentée et soutenue publiquement le 15 décembre 2020

pour l'obtention du titre de Docteur en Sciences Economiques

par

Leonard Sabetti

sous la direction de M. Alexandru Minea et M. Marcel Voia

Membres du Jury

John S. Earle	Professeur à l'Université George Mason	Rapporteur
Angelo Secchi	Professeur à l'Université Paris 1 Panthéon-Sorbonne	Rapporteur
Kim P. Huynh	Banque du Canada	Suffragant
Jean-Louis Combes	Professeur à l'Université Clermont Auvergne	Directeur thèse
Marcel Voia	Professeur à l'Université Orléans	Directeur thèse
Alexandru Minea	Professeur à l'Université Clermont Auvergne	Directeur thèse

Remerciements

In his book *Good to Great*, management guru Jim Collins describes the concept of *the flywheel effect* as a driving force in transformation; a process resembling the relentless pushing of a giant, heavy flywheel, turn upon turn, building momentum until a point of breakthrough, and beyond.

This thesis is the result of many turns of a flywheel. Through graduate training at Indiana University I was exposed to micro-econometrics methods and applications. As a research assistant at the Bank of Canada I learned survey methods and went through a crash course in using Stata. At the World Bank, I was exposed to the innovation literature with a particular focus on measurement and development policy issues. Attending Stata workshops in Ottawa and seminars at George Mason, I learned more about and was inspired by applied research methods. Finally, presenting chapters of my dissertation as works in progress in Clermont-Ferrand and engaging discussants for feedback gave me confidence and encouragement to iron out and make improvements.

Along the way, many individuals accompanied me during this journey for whom I am grateful. First and foremost, I would like to express my deepest gratitude to my supervisors; Marcel-Cristian Voia, Alexandru Minea, and Kim Huynh for their support, guidance and patience. The completion of this thesis is as much their doing as my own. I am also grateful to John Earle and Angelo Secchi for accepting to serve as external thesis examiners and to Jean-Louis Combes for presiding over the PhD jury.

I am thankful to the School of Economics and CERDI, University of Clermont Auvergne for giving me the opportunity to carry out my research towards the dissertation and all the academic team. I am especially grateful to Professor Jean-Louis Combes, Professor Patrick Plane, Nestor Sawadogo and Clément Mathonnat for serving as discussants and commenting on early drafts of my work; Professors Mary-Françoise Renard and Jean-Louis Combes for checking on my annual progress.

I would like to thank my co-author Xavier Cirera for giving me the opportunity to work on innovation policy at the World Bank and absorb learnings from the field and practice of economic development. I am also grateful to Alicia Robb, the Kauffman Foundation and staff at NORC for granting me ongoing access to the dataset on startup firms.

Finally, I would like to dedicate this thesis to my family. Their love and support sustained me over the years.

Sommaire

Introduction générale ……1

Chapitre I. The Returns to Innovation and Imitation in Developing Countries……12

Chapitre II. The Effects of Innovation on Employment in Developing Countries……52

Chapitre III. The Effects of Innovation and Financing on Startup Firm Survival and Growth……84

Chapitre IV. Firm R&D and Knowledge Spillovers……118

Conclusion générale ……143

Bibliographie……148

Introduction Générale

Schumpeter's writing in 1942 highlighted the creative destruction process whereby new innovations make old ones obsolete spurring economic dynamism and transformation. From the post-war reconstruction effort to today's knowledge-driven and global economy, economists have sought to improve our understanding of the growth process from both theoretical and empirical perspectives to better understand its fundamental drivers and inform public policy. At the macro level, Solow (1957) showed how economic growth is driven largely by technological change rather than the accumulation of capital and labor. Subsequently, the emergence of new growth theory emphasized the role of knowledge accumulation in the growth process and the role of the production of ideas as non-rivalrous inputs that generate increasing returns to scale (Romer, 1986).

In contrast to the production of goods and services, non-rivalrous goods create positive spillovers for entrepreneurs to build new products and services on top of an existing idea or alternatively create new ideas that make use of old ones. Some classic examples are the invention of the world wide web or the long list of inventions from Bell Labs throughout the 20th century. However, the production of ideas on their own are insufficient as it takes an entrepreneur with a vision of how to render such ideas useful to solve human problems and create markets – transforming *invention* to *innovation*. Some firms attempt to patent ideas that appear to not have any immediate commercial viability in order to exclude their use by others, which has become a contentious issue in the patent thicket debate, however, even such a practice does not preclude such knowledge from entering the public domain. More often, relying on secrecy, first mover advantage (particularly for platform or scale businesses) and other specialized knowledge internal to the firm provide the needed competitive edge for a time until eroded by competitive forces requiring new rounds of innovation to stay on top. Innovations in organizational, management practices and culture, and not just technological, also play an important role for the success of firms and regions. The ongoing dynamic process of innovation and entrepreneurship has historically played an important role in economic growth.

But where do innovative ideas come from? Peter Drucker wrote that innovation occasionally arises from a flash of genius but more often is the result of a systematic pursuit through hard work and trial and error. Recently, in advanced economies such as the United States business dynamism or turnover of firms has been declining largely driven by a decline in new firms. One potential contributing factor to this decline could be driven by a decline in new

innovative ideas that can be captured by entrepreneurs; Bloom et al. (2020). Major technological breakthroughs seem to only emerge in rapid bursts after many years of aimless, marginal advances. The private sector is often unable to make investments in R&D without some reasonable expectation of a return in the near term and as a result it traditionally occurs in sectors with deep pockets or is financed by mission oriented public sector entities with bold visions. Where will the next wave of innovation and entrepreneurship come from, will there be sufficient spillovers to drive local economies and is there a role for policy? Understanding the wealth of nations is arguably a study on innovation and entrepreneurship: where does it stem from, why does it tend to be clustered in certain regions, is it merely a result of individual human achievement or driven by environmental factors and so forth.

Inspired by these issues, this dissertation aims to contribute to some of our understanding principally through the economist's toolkit. In contrast to the case-study approach, sociological or historical analyses all which provide their own unique perspective, the economics approach typically aims to uncover patterns using observational data towards developing theory about underlying behaviour or outcomes driven by incentives and rational choice, or vice versa. Its focus is on the systematic; not the flash of genius. While any single individual study might struggle to draw convincing conclusions, the success of such an approach occurs over time as a sufficiently large body of evidence is obtained that leads to an improved understanding of underlying causal factors that can inform policy. Our contribution to the literature lies in using new sources of data that make use of improved measures of innovation that help uncover and shed light on underlying drivers of firm performance and heterogeneity. We use data derived from novel firm-level surveys from three separate regions. First, we begin our analysis using data from a large cross-section of developing countries from Africa, South-East Asia, Eastern and Central Europe and the Middle East; these countries are thought to have impressive growth prospects with enormous catch-up potential. A major potential factor in boosting productivity in developing countries stems from the role of innovation in propelling firms towards the technological frontier. Second, the United States, a country viewed to be at the forefront of the technological frontier with an entrepreneurial culture and growth mindset that the rest of the world tends to aspire to and aims to copy. Third, Italy, an advanced economy with both a large small-medium enterprise sector and rich regional diversity that struggles to remain competitive amidst forces of globalization. From a methodological standpoint, we use micro-econometric and program evaluation methods to both incorporate the

nature of the data and to answer questions and untangle relevant margins of interest in a convincing way that are useful for policy evaluation and research while paying close attention to the measurement of economic variables and the role of heterogeneity. Our main goals are both to inform and derive policy implications from the data while contributing to the academic literature and body of evidence and knowledge.

The thesis is situated in the context of a large literature that has sought to understand the sources and dynamics of persistent differences across firms using increasing availability of rich company (or firm) level micro data. Empirical analyses have uncovered how the macro economy is composed of heterogeneous patterns of firm outcomes and performance (including entry and exit). For instance, large mature firms often stagnate after a certain age (Hopenhayn, 1992) and as a result aggregate employment and productivity gains tend to come from entry of fast-growing younger firms (Haltiwanger, Jarmin, and Miranda, 2013). The model of Klette and Kortum (2004) incorporates the concept of innovation from a theoretical perspective that generates rich firm dynamics. Firm innovation and research and development (R&D) strategies have been found to be a major source of explaining differences in firm outcomes (such as productivity) where R&D expenditures are defined as investments to absorb and develop knowledge that can eventually translate to innovations such as new products, processes or other capabilities.

While much of the initial literature focused on studying the effects of competition, market structure and firm size on innovation, more recent research has moved beyond the Schumpeterian debate to focus on determinants of technical advance more broadly such as the role of firm characteristics and industry-level variables such as technological opportunity, appropriability and demand conditions (Cohen, 2010). A primary focus has sought to understand the determination and impact of R&D in terms of its effects on firm performance and patenting. However, limitations of this research and areas for contribution have stemmed from two shortcomings. First, there has been an implicit assumption that all or most innovative activity occurs within formal labs and through R&D accounting expenditures. Second, an absence of accurate measures of innovation itself in terms of its outputs has restricted the analysis to focus on patents as the principle measure. While patents can be an important element of a firm's strategy, they are not necessarily indicative of innovation; particularly for newly launched firms, firms that are far from the technological frontier or firms operating in sectors where patents are not a source of competitive advantage. Even

more challenging from an empirical perspective is that only few firms have any in practice; and when they do, patent valuations have been found to be highly skewed. Any finding of effects on the firm stemming from its patenting activity is likely to be driven by other factors. As a result, improvements in data – such as including more and better data on firms, better measurement of innovative activity, and more granular and richer data on R&D activities as some examples – are enabling a richer understanding of economics of innovation and technological change.

To make progress on the empirical front, incorporating a broader Schumpeterian view of innovation that includes the more incremental implementation of ideas and knowledge to improve the firm in data collection is needed. Under this view, innovation also includes attempts to try out new or improved products or processes and experimentation with alternate ways of doing things. Further, a broader view of innovation could be important for studying firm catch-up, which is likely associated more so with technology adoption, imitation and adaptation; as opposed to a conventional view of innovation narrowly defined as invention, patenting and the use of disruptive technologies. The Oslo Manual guidelines launched an attempt to remedy some of these measurement challenges where the focus is to survey the firm about the quality and degree of novelty of innovation outputs such as product, process and organizational innovations. The Oslo Manual is the basis upon which the OECD and other international organizations collect and publish statistics on business innovation. Increasingly, these enhanced questions on innovation practices are becoming standardized and included in a number of surveys on firms around the world such as the Community Innovation Surveys (CIS) in Europe allowing for benchmarking and comparative analysis (Mairesse and Mohnen, 2010); similar to the pioneering work of Bloom and Van Reenen for management practices.

In these surveys, a formal definition of innovation encompasses "*the implementation of a new or significantly improved product or process, marketing method, or a new organizational method in business practices, workplace organization n, or external relations*" (OECD, 2005). The subcategories of product and process innovation are technological in nature and are the focus of this dissertation; whereas marketing and organization innovation are considered non-technological. Product innovation refers specifically to the introduction of a good or service that is new or significantly improved with respect to its characteristics or intended use. These include significant improvements in technical specifications, components and materials, incorporated

software, user friendliness or other functional characteristics. Further, survey questions differentiate by degree or complexity of novelty; whether the innovation is new to the firm, to the local market or to the international market. In contrast, process innovations are the implementation of a new or significantly improved production or delivery methods. These include significant changes in techniques, equipment and/or software.

Armed with these enhanced measures of innovation outputs, a more complete empirical assessment of the process of innovation at the firm level can be undertaken. The innovation process or function is one characterized by the firm undertaking investments in knowledge activities, both tangible and intangible that include training, purchase of equipment, R&D and licensing that eventually lead to innovation outcomes such as product or process innovations as well as patents. Successful innovations can in turn positively affect firm outcomes such as by increasing productivity, increasing the quality of labor and increasing demand for new products. Further, more productive and innovative firms can also force competitors to exit the market, increasing allocative efficiency. Overall, the effects from innovation at the firm-level will depend on the quality of the innovation introduced as well as the ability of the firm to successfully bring its innovations to market. As a result, a main contribution of innovation stems from both improving efficiency and product quality within the firm and among newly entering firms. In contrast, the economics literature has highlighted the role of reallocation to more productive firms as policy lever to boost per capita incomes; where differences in per capita income have been shown to be roughly 50% explained by differences in firm-level productivity.

The first chapter, *The Returns to Innovation and Imitation in Developing Countries*, begins the analysis in a developing country context using a unique dataset that comprises 53 firm-level innovation surveys from countries in the Africa, South Asia, Europe and Central Asia and Middle East and North Africa regions collected by the Enterprise Survey unit of the World Bank. These data contain representative samples of firms by country that survey their performance amidst their business environment and investment climate and were augmented with an innovation module during the 2013-2015 wave. The innovation survey differentiates between types of technological innovation outcomes such as product and process innovations as well as their degree of novelty. From these data we ask the following set of questions that aim to shed light and fill the gap in the literature on the nature of firm level innovation in developing countries. What are the patterns of

R&D and innovation among firms in developing countries? Does innovation translate into efficiency gains for the firm? Are there differential effects depending on type of innovation?

Obtaining good answers is critical as innovation in developing countries is thought to be less R&D intensive, more focused on incremental adoption of existing technologies or imitation while the returns to such types of innovation are not well understood. Based on studies using OECD country data, the returns to innovation are generally high and thought to be increasing with distance to the technological frontier. Extrapolating to developing countries, one might hypothesize that returns to innovation should be exceptionally large; motivating the prioritization of innovation related investment. From a policy perspective, if returns to innovation are found to be low for existing firms that carry them out, then the potential for lack of enabling factors and complementarities is likely to be high. We rely on the workhorse theoretical model of innovation developed by Crépon, Duguet and Mairesse (1998) that posits the sequential relationship between innovation inputs, innovation outputs and firm productivity. We estimate the model using our data with care to avoid simultaneity and selection bias while incorporating the statistical features of the variables. Overall, our empirical findings reveal an innovation paradox: the coexistence of low levels of innovation-related investment despite the potential for high returns associated with technological adoption. As a result, the promise of Schumpeterian catch-up is far from being realized and there is a need for a deeper understanding of the constraints and impediments to the innovation process.

The Effects of Innovation on Employment in Developing Countries: Evidence from Enterprise Surveys continues the analysis from the first chapter and examines the relationship between innovation and employment. Despite the role of innovation in boosting productivity, there is a tension with the goal of maximizing employment in the short-term; particularly with the rapid wave of digitalization and automation occurring throughout the world. Similarly, reallocation of resources to the most productive or innovative firms in an economy while theoretically sound is inherently in conflict with broader policy objectives of inclusion. Further, the small business sector while typically less productive, plays an important role in creating employment and developing local economies. Fostering innovation throughout the business sector can promote widespread economic gains. As a result, determining whether and to what extent innovation generates a tradeoff between productivity increases and employment growth could be important for

development policy, especially in countries where the needs to absorb new entrants to the labor market in formal and higher productivity jobs are greatest. Are there differences in the impact on employment depending on the type of innovation as well as the firm's distance to the technological frontier? To shed light on these questions using our data, we estimate reduced form parameters from a simple theoretical model that relates changes in production of the firm's new and old products to changes in its labor inputs. Our analysis highlights the role of product innovation as the main channel for employment creation while we do not find any negative impact from process innovation; potentially due to a skill composition effect.

The third chapter, *The Effects of Innovation and Financing on Startup Firm Survival and Growth*, turns to a sample from a cohort of startup firms in the US. Previous literature has emphasized the role of new firms in contributing to business dynamism and economic growth at the macro-level and that particularly only a small share of high-growth startups drives much of these economic gains. At the same time, little is known about the nature of innovation activities among startup firms and how they affect outcomes. To what extent does innovation play a role in the success of the average startup? With this perspective, we study survival and growth of an entering cohort of firms that launch in 2004 and are tracked over an 8-year period until 2011; overlapping with the Great Recession of 2008-2009. Survival and growth are inextricably linked; firms that survive, grow – and vice versa. We study survival to assess firm quality and then subsequently study firm growth as a measure of performance accounting for quality (survival). Undertaking these separate but linked analyses, we study the role and impacts of financing and firm innovation. Financing can affect survival and growth margins but is also endogenous; particularly external financing, as a lender (bank) incorporates its own expectation a firm will survive. We examine the role of financing on the survival rates of firms and whether differential effects are observed as a result of the Great Recession.

While much of the literature on entrepreneurship has focused on the role of financing and the tendency for information asymmetries to result in credit constraints, it is less obvious if entrepreneurship is stifled more so by a lack of financing or rather a lack of ideas. How do measures of firm innovation affect startup firm outcomes? Does financing or innovation matter more for startup firm growth? Hurst and Pugsley (2011) find that many small business entrepreneurs do not want to grow (or modernize). But on the other extreme, some startups are borne out of high-tech

incubator labs, are financed by venture capital funds, and grow and scale very quickly. We hypothesize that the goals and motives of the entrepreneur could matter for untangling firm outcomes and so variables that capture firm innovation activities (whether a firm introduces new products or services along with their degree of novelty) could be important for explaining startup firm growth as well as the ability to raise new funding during the early years of its lifecycle. We contrast the effects from using our measures of direct innovation outputs (product innovation) with traditional measures of R&D and patents that may be less meaningful for new firms where the entrepreneur is likely capturing and exploiting ideas obtained elsewhere. Finally, we compare the effects on firm growth for firms that report innovations versus firms that receive higher levels of financing.

In the final and fourth chapter, *Firm R&D and Knowledge Spillovers*, we study a sample of manufacturing firms from Italy to uncover drivers of firm R&D expenditures. The dataset contains information on firm R&D expenditures, fiscal subsidies for R&D as well as whether the firm cooperates with a university towards R&D. We test hypotheses of crowding-in effects from the university sector and contrast our findings with the impact from fiscal subsidies such as tax credits. Our findings are relevant for innovation policy and maximizing public sector returns on investment.

As private investment in R&D is risky and can lack appropriability, it is thought to be sub-optimal in aggregate and generally undertaken by firms with deep pockets or that operate in sectors defined by monopolistic competition. Among the policy toolkit available to promote innovation, the empirical evidence has highlighted the effectiveness of R&D tax credits and subsidies particularly in the short to medium run (Bloom, Van Reenen and Williams, 2019). However, while such policies can alleviate financial constraints, they are agnostic about underlying R&D project quality. Public sector R&D, including the share that takes place in the university sector (often geared towards basic/fundamental research), has historically been instrumental for unleashing entire new industries and products commercialized by private firms and can also play a role in crowding-in private sector investment through (high-quality) knowledge spillovers. Further, university-firm cooperation on R&D can crowd-in firm-level investment as such partnerships can minimize risk and costs; they also serve as an indirect mechanism to identify and undertake higher

quality investments. Our empirical estimates provide insight into the role of university research and cooperation as a key driver and enabler of private investment in R&D.

As Jean Tirole has written in *Economics for the Common Good*, technological innovation today is at the heart of economic growth and the wealth of nations more so than ever depends on staying on top of the value chain. Many public policy challenges remain around the role of competition policy, intellectual property, culture and institutions, finance and fostering inclusion and the development of human capital; among others. For leading-edge economies, technological innovation can expand the technological frontier generating increases in productivity growth. For economies and firms that are farther from the technological frontier, a major source of catch up growth can occur by acquiring knowledge and ideas and imitating practices and techniques. While countries may be at different stages of development, they all face common challenges. Unleashing the Schumpeterian forces of innovation and entrepreneurship that drive technological innovation could be the key towards sustained economic growth that includes solving humanity's most pressing problems of today. This dissertation aims to paint a picture of innovation and entrepreneurship from an empirical, micro data driven perspective so as to contribute to the body of knowledge that has reinforced our understanding and created impetus for society to embrace entrepreneurship as a powerful force for good.

The Returns to Innovation and Imitation in Developing Countries

Xavier Cirera

The World Bank Group, Washington, D.C.

xcirera@worldbank.org

Leonard Sabetti

Université d'Auvergne, Clermont Ferrand (France)

leonard.sabetti@etu.udamail.fr

The Returns to Innovation and Imitation in Developing Countries

Xavier Cirera*

Leonard Sabetti†

April 9, 2020

Abstract

A major factor in reducing the productivity gap between developing and developed countries stems from the role of innovation in propelling firms towards the technological frontier. This paper uses a unique dataset to empirically characterize the nature of firm level innovation in a sample of 53 developing countries across Africa, South Asia, Eastern and Central Europe and the Middle East and North Africa regions. The analysis shows a negative correlation between the rates of incremental innovation/imitation - defined as innovation new to the firm - and income per capita, which breaks down when more novel/radical - new to the national/international market - forms of innovation are considered. More importantly, when the returns to innovation are estimated, they appear to be positive and high, but only statistically significant in the case of more radical forms of innovation and in regions with higher income per capita. We also document that while the returns to process innovation are low or not statistically significant in most regions, this is likely the result of the definition used in existing surveys. When process innovation is defined by adopting automation processes, the returns are higher than product innovation. These results suggest that many innovative efforts in developing economies do not translate into efficiency gains given the low quality of the innovations introduced and/or a potential absence of key complementary factors.

Key Words: Innovation, Productivity, Developing Countries, Automation.

*We are grateful for helpful comments from seminar participants at the 6th ZEW/MaCCI Conference on the Economics of Innovation and Patenting. We would like to thank CIIP for financial assistance. Corresponding Author: The World Bank, Innovation and Entrepreneurship, Trade and Competitiveness Global Practice, Washington, D.C. Email: xcirera@worldbank.org†

1 Introduction

One of the most important challenges in developing countries is narrowing the existing productivity gap for firms operating far from the technological frontier. Productivity can be understood as revealed firm capabilities (Sutton (2005)) to produce quantity and quality using factors of production more or less efficiently. Innovation that enhances these firm capabilities with the accumulation and investment in knowledge capital combined with management and organizational capital and their tacit knowledge is a critical element for productivity growth. As a result, it is important to better understand the types of innovation activities that are carried out by firms in lower income countries and the impact that these activities have on performance.

However, we know very little about the nature and impact of innovation activities and investments in knowledge capital in developing countries. The view that innovation is the work of highly educated skilled labor in research and development (R&D) intensive companies with strong ties to the scientific community, is inevitably more a first world perspective than an appropriate description of innovation activities in developing countries (Fagerberg, Mowery, and Nelson (2006)). Innovation in these countries appears to be less R&D intensive and with minimal patenting. In these countries, innovation is better described as attempts to improve and imitate new products, processes and adopt existing technologies (Bell and Pavitt (1997); Kline and Rosenberg (1986)). This is a process of technology adoption, imitation and adaptation far from the technological frontier, with little formal R&D, where firms adopt incremental (as opposed to radical) changes (Fagerberg, Mowery, and Nelson (2006)).

Given this different pattern of innovation activities, the critical question becomes whether different types of innovative efforts have an impact on performance and on reducing the firm's distance to the technological frontier. This is especially important given that the riskiness associated with introducing innovations is potentially higher for firms in developing countries; where the lack of complementary factors to innovation, market failures, policy distortions and investment climate costs are typically high. From a policy perspective, in order to maximize the contribution of innovation on firm-level productivity growth, we need to determine whether firms that innovate experience sufficient returns to their investments. If the returns are low, firms will be reluctant to invest in enhancing their innovation capabilities. While, Griffith, Huergo, Mairesse, and Peters (2006) suggest that the returns to R&D may be greatest for countries farther away from the technological frontier for a sample of OECD countries, Goni and Maloney (2014) find for a sample of developing countries that the returns to R&D follow an inverted U shape - low for poorer countries and increasing only until a point where the lack of

complementary factors lead to diminishing returns to R&D. A similar dynamic occurs with the introduction of innovation outcomes, such as product and process innovation. Most of the existing evidence in OECD countries finds a positive impact of innovation, especially product innovation, on productivity (See survey by Mohnen and Hall (2013). However, the evidence for developing countries is scant and mixed. Some exceptions are Crespi, Arias-Ortiz, Tacsir, Vargas, and Zuiga (2014) for Latin America who find positive returns, while Goedhuys, Janz, and Mohnen (2008) for Tanzania and Cirera (2015) for Kenya find no statistically significant impact of innovation outcomes on productivity.

In addition to lack of complementary factors, it is important to highlight the importance of firms' capabilities to absorb knowledge in developing countries. Ideas that generate innovation are not costless, and information even when freely available requires some investment or capability on the part of the user in order to fully appropriate its benefits. For example, Cohen and Levinthal (1989) describe internal R&D efforts of firms as a dual process of creating new knowledge but also enhancing their ability to assimilate and exploit external knowledge, or "absorptive capacity". This "absorptive capacity" is likely to be low in developing countries given the significant shortage of skills and low quality of managerial practices.

This paper aims to fill this gap in our understanding of the nature of firm level innovation and the link to productivity in developing countries. Specifically, the objective of the paper is threefold. The first objective is to provide a description of the nature of firm level innovation activities in developing countries. The second objective is to measure the returns to innovation in developing countries. The final objective is to decompose how these returns (if any) vary by region, sector and the degree of novelty of the innovation introduced. In order to do so, we use a unique dataset that comprises 53 firm-level innovation surveys from countries in the Africa, South Asia, Europe and Central Asia and Middle East and North Africa regions collected by the enterprise survey unit of the World Bank in the period 2013-2015. To our knowledge this is the most comprehensive and comparable cross-country dataset with innovation information.

To advance some of main conclusions, the analysis confirms some of the findings of the case study literature and suggests differences in the nature and impact of innovation in these countries. First innovation efforts are primarily non-R&D based. Second, there appears to be a negative correlation between the rates of incremental innovation/imitation - defined as innovation new to the firm - and income per capita. This correlation, however, disappears when more novel/radical - new to the national/international market - forms of innovation are considered. Third, the returns to innovation are positive and high, but only statistically significant in the case of more radical forms of innovation

and in regions with higher income per capita. Forth, the returns to process innovation are low or not statistically significant in most regions; which is likely the result of the definition used in existing surveys. When process innovation is defined by adopting automation processes the returns are higher than product innovation.

The paper is organized as follows. Section 2 provides a brief review and summary of the literature and empirical evidence on the impact of innovation on productivity. Section 3 discusses our empirical methodology and section 4 describes the data used in our analysis. Section 5 presents our results while Section 6 concludes.

2 Innovation and Productivity. A Brief Review of the Literature.

A dominant Schumpeterian view of the growth process sees innovation as the engine of the creative destruction process that spurs economic dynamism and transformation. This view started to permeate growth theory since the seminal work of Solow (1957), which placed innovation at the center of a growth process driven by technical change. This influence was reinforced with the emergence of new growth theory emphasizing the role of knowledge accumulation for the growth process and Schumpeterian creative destruction arising from a competitive R&D sector as the main engine of growth (Aghion and Howitt (1992); Romer (1986)).

This increasingly central role of innovation in endogenous growth theory was not accompanied by a similar effort to formalize the rich micro dynamics of firm level innovation that underpin the growth process. Innovation is primarily a micro phenomenon. It is the outcome of firms' investments in knowledge capital and management decisions. The ultimate objective of these investments is to introduce improvements and new products and processes that positively impact firm performance by increasing productivity, employment, sales, profits, market share or markups. However, there is uncertainty regarding the extent to which firms are able to convert knowledge capital investments into innovation outcomes and furthermore, whether these innovation outcomes are likely to impact firm performance. Innovation is risky since it is almost impossible to determine ex ante whether the introduction of new products, processes or organizational changes will lead to an increase in sales, employment or productivity; and also, whether these firm outcomes will impact the reallocation of factors and firms' entry and exit.

In the last decade, however, there has been several attempts to model these rich micro firm dynamics arising from innovation efforts in the tradition of the Schumpeterian creative destruction. Klette and

Kortum (2004) develop a general framework of reallocation where firm level innovation increases product quality and make firms more competitive, which increases their revenue and size; and forces existing firms producing old and obsolete versions of the product to exit the market. Without these innovation forces in action, models of firm dynamics typically predict that firms stagnate after reaching a certain size, conditional on survival, as in Hopenhayn (1992). Innovation allows in these heterogeneous firm models the churning of firms that is widely observed in the data.

Lentz and Mortensen (2008) extend the framework by Klette and Kortum (2004) to include firm heterogeneity. This allows to replicate the observed large dispersion in productivity levels across firms. In the model, innovation investments increase product quality and allow firms to capture higher prices and profits, which allow them to become more productive and larger. Akcigit and Kerr (2010), and Acemoglu, Akcigit, Bloom, and Kerr (2013) develop heterogenous firm models that also allow exploring reallocation of resources between entrants and incumbents driven by innovation efforts. In Acemoglu, Akcigit, Bloom, and Kerr (2013) firms defer on their capacity to transform innovation on productivity growth, which creates reallocation of skilled labor across types, and dispersion and exit of low capacity entrants and incumbents. A common theme of these models is the importance of the creative destruction process and the central role of innovation driving it.

These models have been calibrated and their predictions validated using US firm level data. An important question, however, is whether this process of creative destruction is also in motion in developing countries, given that barriers to entry and distortions are likely to be more prevalent in lower income countries, see Hsieh and Klenow (2009), and act as a constraint for the reallocation of resources. A first step, however, to analyze these micro dynamics is related to the ability of firms in developing countries to produce innovations, the nature of these innovations and the ability to convert these in productivity growth.

A large literature has examined the relationship between innovation and productivity at the micro level. Most of the evidence, however, has used firm level datasets in OECD countries. Hall (2011) provides a comprehensive survey of the empirical work in this area. The survey focuses mainly on 16 existing micro empirical studies using the workhorse empirical model of innovation, the Crepon-Douget-Mairesse (CDM) model (Crepon, Duguet, and Mairesse (1998)). The CDM models the relationship between innovation and productivity as a sequential process in three stages. In the first stage, the determinants of knowledge inputs are estimated - R&D and potentially other inputs. In the second stage, the relationship between innovation inputs and innovation outcomes - product and/or process - is determined. Finally, the last stage estimates the impact of those innovation outputs have on productivity

- proxied by sales per worker. Hall (2011) main finding is that in general most studies find a positive correlation between product innovation and productivity.

One important regularity in the findings for developed countries, although slightly counter-intuitive, is the fact that product innovation tends to have a larger impact on firm-level productivity than process innovation, and even some studies find a puzzling negative impact of process innovation on productivity. According to Hall (2011), the problem with process innovation is that it cannot be measured in the surveys beyond the dichotomous variable of whether or not the firm implements process innovations. More importantly, this variable encapsulates all types of processes, ranging from delivery methods to automation.

One final result in the empirical literature is the heterogeneity in the size of the impact across sectors. For example, Criscuolo (2009) show that the productivity elasticity of innovation tends to be higher in manufacturing than in services sectors. Mohnen and Hall (2013) confirm these results reviewing a larger sample of studies. Overall, the evidence gathered by these empirical studies suggests that the introduction of innovations outcomes by firms, primarily new or upgraded products, have a positive impact on productivity.

The existing evidence for developing countries is, however, scarce; especially regarding low-income countries. One relevant study is Goedhuys, Janz, and Mohnen (2008), which examines the main drivers of productivity in Tanzania. The authors do not find any link between R&D, product and process innovations, licensing of technology or training of employees and productivity. The results suggest that Tanzanian firms are struggling to convert knowledge inputs into productivity improvements due to the poor enabling environment for business, which is the main constraint on productivity according to their empirical results. In addition, there are indications that in developing countries, investments in knowledge capital are smaller than in developed countries. Goni and Maloney (2014) show that investments in R&D are smaller in developing countries as a share of GDP than in developed countries. As suggested above, one potential explanation for this result is the absence of complementary factors that enable an efficient use of R&D; such as education, the quality of scientific infrastructure, and the private sector; which is weaker in countries far away from the technological frontier. One implication of this knowledge capital deficit is the possibility that the quality of the innovations produced and, therefore, the returns to innovation are lower in developing countries, see Cirera, Iacovone, and Sabetti (2015).

A related strand of the literature has empirically analyzed some of the complementarities needed in innovation. The literature has mainly focused on complementarities between innovation outputs;

modes of innovation. The idea is that introducing different types of innovation outputs jointly have an additional impact on productivity. Polder, Van Leeuwen, Mohnen, and Raymond (2010) find significant complementarities between different knowledge inputs and innovation outcomes, in the Netherlands. The authors find that (i) ICT investment and usage are important drivers of innovation; (ii) a positive effect on productivity of product and process innovation when combined with organizational innovation; and (iii) evidence that organizational innovation is complementary to process innovation. Miravete and Pernas (2006) find evidence of complementarity between product and process innovation in Spanish tile industry. There is also, some evidence of complementarities of knowledge inputs. For example, Cassiman and Veugelers (2006) find important complementarities between internal R&D and external knowledge acquisitions for the introduction of innovation outcomes.

A final element especially important for developing countries when considering the relationship between innovation and productivity in developing countries is the lack of an enabling business environment. Barriers to physical accumulation are also likely to undermine investments in knowledge accumulation. Also, existing market failures in the supply of technical infrastructure, human capital or technology are also likely to reduce the quality of the innovation inputs introduced and the returns to innovation.

3 Methodology

3.1 Empirical strategy

In order to estimate the relationship between innovation and productivity given the nature of our data, we use the CDM model Crepon, Duguet, and Mairesse (1998) briefly described below. Innovation requires the transformation of knowledge capital or innovation inputs, both tangible and intangible, such as training, equipment, R&D or the acquisition of intellectual property into innovation outcomes- the introduction of new products, improved quality, new production processes or organizational changes. Firms invest in accumulating these knowledge inputs in order to increase their capabilities and produce innovative outcomes. Once innovation outcomes have been introduced by the firm, these can impact firm performance in different ways. Successful innovations are likely to increase firm-level productivity by improving the capacity to transform factors of production into new and/or more efficiently products of higher value or quality. Second, the increase in productivity is expected to increase the marginal productivity of labor, and as a result, increase the quality of jobs (i.e.: more productive jobs). Third, more productive firms are expected to push less productive firms out of the

market, increasing allocative efficiency, thereby increasing the overall efficiency of the economy. All this, however, depends on the quality of the innovation introduced and the ability of firms to successfully market new products and sustain new processes.

The CDM model by Crepon, Duguet, and Mairesse (1998) can capture most of the features described above. The model contains a recursive system of equations that characterize the evolution of the innovative process in three stages. An initial set of equations in the first stage explain the firm's decision to invest in R&D, sometimes referred to as the knowledge equation, influenced by firm specific, regional and other factors. The second stage characterizes innovation outcomes, or the likelihood that the firm's investment in knowledge capital translate into successful innovations. In the third stage, an equation links innovation outputs to firm productivity.

Solving each of the equations of the model separately raises endogeneity issues, which would otherwise lead to biased coefficients and underestimating the true effects of innovation on productivity. For example, unobserved factors that make a firm more likely to invest in R&D, such as unobserved managerial ability, might be correlated with firm's higher levels of productivity (selection). Alternatively, unexpected shocks that lead to higher chances that innovation outcomes are successful may also (positively) affect productivity outcomes in the performance equation (simultaneity). Consequently, solving the system simultaneously as in the CDM framework can correct some of these issues.

The literature has tended to use two different empirical strategies when looking at the relationship between innovation and productivity. One approach pioneered in the original CDM model (Crepon, Duguet, and Mairesse (1998)) uses Asymptotic-Least-Squares, which embodies the jointly estimation of the main equations of the model. The second approach, see Griffith, Huergo, Mairesse, and Peters (2006), estimates a sequential model, using IV in each stage to account for endogeneity. Under this approach, predicted values of endogenous variables in the first stage are included in the estimation of the second- and third-stage equations, and correcting the standard errors by bootstrapping methods. Hall, Lotti, and Mairesse (2009) and Musolesi and Huiban (2010) compare the results from both methods and find that there are not significant differences in the impact of innovation on productivity as long as endogeneity and selection are properly treated. Hall, Lotti, and Mairesse (2009) find similar results comparing both types of estimation, although larger standard errors using simultaneous maximum likelihood of all three stages than the sequential method. Mohnen and Hall (2013) discuss the tradeoffs between efficiency and robustness to misspecification.

Recent advances in computation, however, have opened the possibility of using more efficient estimation methods. Specifically, Roodman (2011) shows that the case of seemingly unrelated equation

(SURE) systems that falls into a class of mixed processes can be modeled recursively where each stage is fully observed. Our model falls into this category and therefore a full information maximum likelihood (FIML) estimator can obtain consistent estimates for the model's endogenous, right-hand side parameters. Our Structural Equation Modelling (SEM) approach to estimating the CDM model is similar to Baum, Loof, Nabavi, and Stephan (2015), who estimate a generalized structural equation model of the R&D, innovation and productivity relationship using a panel of Swedish firms from the Community Innovation Survey; and Schfer, Stephan, and Mosquera (2015) who study the role of financing constraints on innovation outcomes and productivity using the 2007 Mannheim innovation survey.

In the original CDM model, the authors solved the model simultaneously using Asymptotic Least Squares (ALS), a minimum distance estimator.[1] This approach is advantageous when the joint distribution of observed variables does not have a closed form and when numerical integration techniques are computationally burdensome. Consistent estimates are first obtained in each equation via maximum likelihood, and then the structural parameters are obtained by minimizing their distance to the reduced form estimates using the empirical covariance matrix weighting function. One benefit to this approach lies in its ability to take into account measurement issues arising in the nature of survey data. For example, firms that do not report undertaking R&D may in fact still possess knowledge outputs. However, one drawback from this approach is its recursive structure, which in principal does not allow for feedback effects. Specifically, each block of equations representing a respective stage of the innovation process is nested in the subsequent stage. Furthermore, standard errors may be underestimated in the model's last stage.

The SEM methodology implemented on observational data can be interpreted analogously to the experimental or program evaluation approach that seeks to obtain causality or identification from natural or quasi-natural experiments. Under this setting, when an intervention is completely random, the difference between the mean outcome from the control group and a treatment group form an average treatment effect. However, usually a selection on observables assumption is required and treatment effect estimates are susceptible to bias in the presence of unobserved heterogeneity. Under the SEM framework, identification depends on the quality of exclusion restrictions, functional forms and joint normality assumptions. See Wooldridge (2015) for a discussion on the use of control functions for solving the problem of endogenous explanatory variables in non-linear models. One key advantage of our approach is the mixed nature of the estimation, which allows to accommodate non-linear processes

[1]Alternatively, the model could be solved by GMM but it has been shown to be less efficient in large samples.

and fully exploit the complex interactions in the data without relying on approximations to functional forms. This is critical given the censored and dichotomous nature of some of the data.

3.2 Overview of the Model

1.1.1 The Knowledge Function

The first step of the model is to specify the choice of knowledge capital investment intensity. To measure knowledge intensity, we use the firm's investment in R&D per worker. Although, there are other important knowledge activities that firm carry out to innovate such as training or use of intellectual property, the only information available for all the countries in our sample is R&D. It is likely that some of these investments are correlated with R&D intensity and, therefore, the coefficient on R&D captures not only the effect of R&D on innovation but also the impact of other knowledge investments. Also, as shown in the next section, a large number of innovative firms do not use any formal R&D. Unlike in Griffith et al. (2006) we have information on R&D investments for all firms, including non- innovators, in the enterprise surveys. As a result, we take into account the fact that the knowledge capital investments of some firms do not translate into innovation outcomes, as well as the fact that those investments are zero for many firms. One challenge with R&D information in our data is that firms find it very difficult to accurately respond the amount of investments on R&D. As a result, we use a probit model to include on the model the decision to invest in R&D. In the sensitivity analysis we also try to model the intensity of R&D. Specifically, we estimate the following model

$$R\&D_i = X_{i1} + \varrho_{i1}, \tag{1}$$

We characterize the determinants of investing in R&D by three sets of factors that are commonly found in the literature (in the Schumpeterian tradition): 1) firm characteristics, 2) the firms market conditions (including demand pull and technology push factors), and 3) investment climate conditions (See Table 3). Firm characteristics include size, age, the presence of foreign capital ownership, the degree to which the firm is diversified or operates in multiple markets, and industry. We also include measures of the degree of the firm's capital intensity measured by the capital per worker ratio as well as the share of skilled employees with formal education. The firms market conditions can be important for incentivizing investments in innovation inputs. Demand pull factors are captured by changes in the firm's market share over the past three years, the degree of integration in international markets such as whether the firm exports or imports, and whether the firm competes with the informal sector and the extent to which informal firms pose a major obstacle to the firm's competitiveness. Technology push factors

can be expressed either as the firm's perception of how important these factors were in influencing its innovation activities or can be based on actual reported measures that serve as proxies. We use whether the firm has obtained certification and whether it has licensed any foreign technology. The investment climate can be an important determinant in firm R&D as a lack of adequate assurance that a firm may reap benefits from its investments in knowledge inputs may deter these efforts in the first place. Our dataset contains information on the extent to which the firm lacks access to finance, various infrastructure related obstacles in trade, telecommunications, educated workforce, and the regulatory environment.

1.1.2 The Innovation Function

We use a Probit model to estimate the probability the firm undertakes innovation, controlling for a set of determinants that include the firm's investment in knowledge capital. Another important input into the innovation function is the number of technical staff in the establishment that can facilitate the transformation of knowledge inputs into innovation outcomes. Given that the data on skilled labor is not available in all countries, we proxy skilled labor by an index of the extent to which an inadequate labor force poses an obstacle.

$$i_i^* = \beta_3 X_i + \gamma_1 R\&D_i + \varrho_{i,2}. \tag{2}$$

1.1.3 Productivity Equation

Productivity is derived from a Cobb-Douglas function where sales (Y) are a function of capital (K), labor (L) and augmented with innovation outputs (I).

$$\begin{gathered} Y = f(I, K, L), \\ Y_i = I_i K_i^{\alpha} L_i^{\beta}, \\ \frac{Y_i}{L_i} = I_i \frac{K_i^{\alpha}}{L_i^{\alpha}} \frac{L_i^{\beta}}{L_i^{\beta}}, \\ where\ \alpha + \beta = 1\,. \end{gathered} \tag{3}$$

We transform the equation 3 logarithm form, add sector controls, X_i, and augment it with an indicator variable for innovation.

1.1.4 Mixed processes model

Combining equations, we have the multivariate system:

$$
\begin{gathered}
log\left(\frac{Y_i}{L_i}\right) = \delta_0 + \alpha log\left(\frac{K_i}{L_i}\right) + \beta_{innov}(i_i^* > 0) + \delta_1 X_i + \epsilon_i, \\
i_i^* = \beta X_{1_i} + \gamma_1 R\&D_i + \gamma_2 Z_i + \eta_i \\
R\&D_i = X_{3_i}\beta + \nu_{2_i}. \\
\epsilon_j, \nu_j and\ \eta_j \sim MVN(0, \sigma),
\end{gathered}
\tag{4}
$$

where the error terms of the equation are distributed as multivariate normal with mean zero and variance Σ. This system of equation is estimated using Geweke–Hajivassiliou–Keane (GHK) simulated maximum likelihood and is implemented in Stata; see Drukker (2011) for more details.

4 A Profile of Firm Level Innovation in Developing Countries

4.1 The dataset

In order to study the role of innovation on firm performance in developing countries, we use the World Bank Enterprise Survey and its linked innovation modules that were implemented during the period 2013 to 2015. This is the most comprehensive set of cross-country surveys on innovation information carried out to date using an almost identical questionnaire. The survey used a stratified sampling strategy, where firms in the formal sector are stratified by industry, size and location.[2] Since in most economies the majority of firms are small and medium-sized, Enterprise Surveys tend to over-sample larger firms. This is also the case in most innovation surveys around the world.

An additional advantage of the survey is that it collects substantial balance sheet and other information regarding the investment climate, which enables the linkage of innovation efforts to performance and potential obstacles. The core questionnaire of the Enterprise Survey includes information in a large array of issues: firm characteristics, gender participation, access to finance, annual sales, costs of inputs/labor, workforce composition, bribery, licensing, infrastructure, trade, crime, competition, capacity utilization, land and permits, taxation, informality, business-government relations and other performance measures. In addition to the core questionnaire, an innovation survey was implemented separately to 75% of the firms in the sample. The core innovation questionnaire is based on the Oslo manual guidelines[3] and the questions are very similar to the Community Innovation Surveys (CIS) questionnaires implemented in the European Union. The mode of data collection in both surveys is face-to-face interviews.

[2] Firm size levels are 5-19 (small), 20-99 (medium), and 100+ employees (large-sized firms)
[3] Available here: http://www.oecd.org/sti/inno/oslomanualguidelinesforcollectingandinterpretinginnovationdata3rdedition.htm

Table 1 outlines the list of countries in our sample as well as the number of firms that completed the innovation surveys. In total, the pooled sample with innovation questions is 33,462 firms for 53 countries. Most firms in the sample are concentrated in the manufacturing sector (49%) and wholesale and retail (29%); although the sector composition varies by country. The countries with the largest representation in the sample are Russia (12%) and India (10%), and some countries in ECA have less than 250 firms in the innovation modules. In some countries in the EU, however, the sample size appears to be very small for their economic size; for example, Slovenia, Czech Republic and Hungary. Table 2 provides the size and sector decomposition for the pooled sample. As suggested above, only formal firms are surveyed.

4.2 Innovation profile

A first empirical fact arising from the data is the confirmation that most innovation occurs without formal R&D activities; sometimes even for more radical innovation defined as new to the national or to the international market (see Figure 1). On average, only roughly 20 percent of firms conduct R&D based innovation, but this share roughly doubles when only more radical innovation as defined by new to the international market is considered. When R&D is low, innovation is likely to be less novel. We also observe in the sample very low firm appropriation patenting activities, which may be partly associated to the large prevalence of imitation in innovation - new to the firm or to the local market and low R&D activities; as predicted by Schumpeterian theories where intellectual property rights incentivize investments in innovation and enable inventors to reap the rewards of their efforts and risk taking.

Figure 2 plots R&D intensity rates measured as average firm expenditures per worker by GDP per capita, a proxy for distance to the technological frontier. A clear positive correlation exists as firms in countries closer to the technological frontier tend to be more R&D intensive. Cohen and Levinthal (1989) describe internal R&D efforts of firms as a dual process of creating new knowledge but also enhancing their ability to assimilate and exploit external knowledge, or absorptive capacity, which is weaker in firms farther away from the technological frontier. The bulk of R&D is intramural; extra-mural R&D is marginal. While poorer countries tend to do less R&D, R&D expenditure tends to be concentrated in a few firms and on average R&D intensity is very low in most countries, especially in Africa and South Asia. Firms that are larger, participate in international markets, in high-tech manufacturing or knowledge intensive services sectors have higher prevalence of investing in R&D. In relation to other types of knowledge investments, the purchase of equipment is the main other source

of knowledge acquisition; while the purchase of intellectual property is negligible. However, data on other types of knowledge activities such as equipment purchase, training and acquisition of knowledge such as licensing, patents and inventions, is only available for countries in Africa and South Asia and hence our measure of R&D investment in our econometric estimates excludes other categories of knowledge.

Looking more specifically at innovation outcomes, the innovation survey differentiates between two technological innovation outcomes, product and process, and two non-technological innovations, marketing and organization. However, there are some cognitive problems with the way innovation questions are formulated.[4] The most important problem is that there is significant confusion when identifying the different types of innovation outcomes by firms in the survey. For example, new marketing processes such as discounts, new packaging or new client segments are sometimes misunderstood as process or product innovations. The fact that interviewees provide a recorded description of the product and process innovations that they claim they have introduced allows us to verify these innovations, reclassify wrongly attributed cases to their respective category or invalidate cases that do not constitute an innovation at all (the detailed methodology to clean the data is described in Appendix 2). Overall, the cleaning exercise results in a decrease in both product and process innovation rates due to cases of incorrect classifications of innovation or misclassifications of marketing as product or process innovations. Figure A1 in the Appendix shows the before and after cleaning country averages of product and process innovation. The reduction in innovation rates due to cleaning is significantly large in the Africa region and Eastern Europe, and especially for product innovation, and suggests that using raw data from innovation outcomes is potentially leading to biased interpretations given its subjective nature and the lack of clarity around the definition of the different types of innovation. In order to mitigate some of these biases, in what follows only the cleaned data on innovation outcomes are used.

Overall, innovation rates using the baseline definition of new to the firm, are significantly high and interestingly tend to be negatively correlated with GDP per capita. Product innovation rates are on average 26 percent in Africa, 29 percent in South Asia, 19 percent in ECA and 16 percent in MENA. Process innovation rates tend to be higher in Africa (31 percent) and South Asia (35 percent), compared to ECA (13 percent) and MENA (12 percent). As shown in Table 4, firm-level innovation rates

[4]See Cirera and Muzi (2016) for a description of the data and the problems with this type of survey when using innovation questionnaires.

fall significantly, especially in Africa and South Asia, when we depart from the definition of at least new to the firm (imitation) to more stringent degrees of novelty based on whether the innovation is new to the national and international markets (radical innovators). However, radical innovation appears to have little correlation with income per capita as shown in Figure 3. This suggests that the bulk of the innovation activity, especially in these two regions, corresponds to firm incremental innovation or pure firm imitation as has been advanced in the innovation literature.

These average regional rates described above mask significant heterogeneity across countries. Within the same region, the ratio of innovation rates - considering both product and/or process innovation between high and low innovative countries is in some cases tenfold. For seven countries; India, Bangladesh, Namibia, Zambia, South Sudan, Kosovo and Uganda, more than half of the firms have introduced product or process innovation. Symmetrically, in six countries; Georgia, Albania, Turkey, Sudan, Uzbekistan and Azerbaijan, less than 10 percent of firms have introduced either a product or a process innovation or both. Therefore, there appears to be very different innovative dynamics within and between regions. In general, larger firms tend to display higher innovation rates, irrespective of the type of innovation, although the pattern is not always monotonic, and the gap between small and large firms is not very large.

While higher levels of imitation and technology adoption should be expected in developing countries, an important question is whether we observe positive returns from this less novel and more incremental type of innovation, since marginal improvements may have little impact on efficiency. This may be especially the case for process innovation, since the same definition includes any process upgrade ranging from marginal changes in production processes to automation, which can have very diverse impacts on efficiency. In the next section we attempt to measure the impact of innovation on productivity in developing countries, decomposing the results according to the novelty of the innovation introduced and the type of process innovation.

5 Results

We first estimate the baseline structural model described in section 3, which follows closely the structure of the CDM model. The model is estimated on the overall sample using sector (ISIC 2 digits) and country dummies (Table 5) using FIML. The variables used are based on the literature introduced in Section 3 and described in Table 3; and correspond to firm characteristics, market structure and access to

finance, technology push factors and spillovers.

Given the large heterogeneity in innovation outcomes across firms we estimate four different specifications trying to capture two dimensions of innovation outcomes: product vs process innovation, and the degree of novelty of the innovation outcomes introduced. In specifications (1) and (2) we separate between product and process innovation, while in specifications (3) and (4) we use a unique output indicator for technological innovations - product and/or process. In specifications (2) and (4) we further split innovations based on whether the innovation introduced is new to the firm or the local market (imitation) or whether the innovation is new to the national or international market (radical). Given the fact that the extent of these innovations is qualitatively different we specify different processes for both types of innovation and, therefore, also add to the model imitation and radical product innovation equations in the second stage. In what follows we describe the results stage by stage of the model.

5.1 Determinants of R&D incidence

The bottom of Table 5 shows the marginal effects for the first stage, R&D incidence. In line with theoretical predictions, larger firms are more likely to engage in R&D activities; although we do not find any impact of age on the incidence of doing R&D. More importantly, both importers and exporters are also more likely to engage in R&D investments, which emphasize the role of participation in international markets in investing in knowledge activities. This can be the result of access to better knowledge, inputs and technology through international trading, and through more sophisticated demand in international markets forcing firms to become more competitive and increase their knowledge investments.

Interestingly, we do not find any significant impacts from our market structure variables whether the firm operates in a monopoly/duopoly or the extent to which competition from the informal sector is an obstacle, since it is likely that firms doing R&D do not compete directly with informal sector firms. One explanation of this result is the low quality of these proxies of market structure, given the difficulties to build firm specific variables of market competition that are not endogenous to innovation and productivity, such as changes in market shares and markups.

Existing technology factors such as updated capital stock and license of foreign technologies are important correlates of engaging in R&D, since these tend to be complementary factors to R&D activity. More importantly, lack of access to finance proxied by the share of working capital financed by own resources appears as an important deterrent of R&D activities, suggesting that the innovation

financing gap may be playing an important role in constraining R&D and the quality of innovation outcomes.

5.2 Determinants of innovation outcomes

In the second stage we estimate the probability of introducing innovation outputs. We differentiate between product and processes in some specifications while in others we merge product and process innovation in one indicator. In general, and as expected, firms that perform R&D are more likely to innovate. Interestingly, when we estimate the effect of R&D on new to the firm or to the local market - imitation - the effect is marginally or not statistically significant; suggesting that R&D is not an important input for the less novel types of innovation that only require some form of product imitation.

Similarly, to the case of the determinants of R&D, larger firms and those participating in international markets via exporting and importing are more likely to innovate but more so when more radical forms of innovation are introduced. Again, the impact of age is not a significant predictor for most forms of innovation. Education as an obstacle as well as demand pull effects tend to positively affect innovation in all cases. In the case of skills shortage, it is likely that only those firms that try to introduce innovation are the ones that perceive to be more constrained, explaining this positive coefficient. The impact of the duopoly/monopoly is mainly not statistically significant, only affecting negatively some forms of technological innovations. Again, lack of access to finance proxied as the share of working capital that is self-financed affects negatively the probability of innovating.

We also introduced agglomeration and spillovers proxies in the decision to innovate. For agglomeration we use a dummy for whether the firm is located in the main business city. The estimate for this variable is mainly positive and statistically significant when more radical forms of innovation are considered. Regarding, spillover effects, we use the share of innovators within the same region. Spillovers matter for all types of innovation. The role of knowledge spillovers has been highlighted in the literature since Jaffe, Trajtenberg, and Henderson (1993).

Overall, the specifications used based on the innovation literature tend to have a larger explanatory power for more radical innovations, which reinforces the idea that there is a continuum of innovation possibilities according to the distance to the technological frontier, and that different types of innovation require different inputs and R&D is less important at the imitation side of this spectrum.

5.3 The impact of innovation on productivity

Given the significantly high rates of innovation, especially in Africa and South Asia, an important question is whether these translate to significant returns. In addition, one stylized fact of the productivity literature is the fact that productivity is highly disperse even within firms in the same sector (see Syverson (2011)), especially in developing countries (Hsieh and Klenow (2009)). As a result, innovation at the firm level, which typically has been an unobserved factor, may be accounting for some of this productivity heterogeneity. Innovation might play a dual role in terms of enhancing the productivity distribution of firms in an economy. On one hand, marginal innovations may lead to increases in productivity for firms in the lower end of the distribution; for example, these firms may be using older technologies, have yet to update their IT infrastructure and so forth, or are providing products to the market that are marginally less cost effective. On the other hand, radical innovations have the potential to be productivity enhancing for the firm that introduces them, but also may have aggregate wide effects by rendering inefficient firms obsolete, and at the same time the creation of new demand has the potential to generate spillover effects where new industries emerge to provide downstream or complementary products and services. These effects are also likely to contribute to the dispersion of productivity.

The last stage in the top of table 5 shows the results for the productivity equation. Overall, we find that the returns to innovation in developing countries are positive and high; and in line with the literature the returns are larger for product than for process innovation. The estimates of the productivity stage of the structural model pooling all the sample and controlling for sector and country effects suggest a return to innovation that are 213% $(100 * (\exp(1.143) - 1)$ for product innovation and 24% for process innovation $(100 * (\exp(0.214) - 1)$. When considered together, product and process innovation, in specification (3) the returns are 134%.

Specifications (2) and (4) decompose innovation according to the novelty of product innovation - imitation when new to the firm and local market vs radical when new to the national or international market. As expected, the results show that although positive and statistically significant the coefficient for imitation is between 3 and 5 times smaller than for radical innovation; which is equivalent to 4 and 6 times larger in percentage terms. Both types of innovations matter, but the impact of radical innovation is significantly larger.

These results are somewhat similar to the existing evidence for OECD countries in sign but not in magnitude. Compared to studies using similar specifications summarized in Mohnen and Hall (2013) our pooled estimates are significantly larger, to a magnitude of ten times larger in comparison with some

OECD countries, especially when we disaggregate between product and process innovation. Although the literature is not unanimous on what type of product innovation has larger returns, a significant number of studies have found higher returns to product innovation.

5.3.1 Disaggregating process innovation: the impact of automation

The fact that the returns to innovation are larger for product than for process innovation is somehow puzzling. On the one hand, it is possible that the fact that we proxy productivity as sales per worker could play a role explaining this result, since new or upgraded products may increase market power and, therefore, sales per worker; reflecting more changes in market structure than real efficiency gains. On the other hand, one would expect that the largest productivity effects are related to the introduction of more efficient technologies - process innovation.

The current measuring framework for innovation, the Oslo manual, includes, however, a different set of processes under process innovation, such as delivery methods and other processes used in the production of goods. The question, therefore, is to what extent the results would change when using as process innovation only those cases where the innovation introduced imply to some extent automation of the production processes. To explore this question and taking advantage of the fact that our survey includes an explanation of the process innovation introduced, we construct a new indicator variable that takes value one when the description of the innovation suggests that the new process included some degree of automation. Figure A2 in the Appendix shows the percentage of process innovation firms where it involved some degree of automation of the production process.

Table 6 shows the results equivalent to the benchmark model in Table 5 but replacing process innovation by automation. The results are striking: the coefficient for process innovation increases six-fold, increasing the returns from 24% to 250%. The coefficient on technological innovation either product and/or automation also increase. In some of the specifications, process innovation moves to be statistically significantly larger than product innovation, including imitation and radical innovation. This suggests that process innovation can have large productivity effects, but often is poorly measured or includes other types of process changes that are not necessarily more efficient.

5.3.2 Decomposing the productivity effect by income, region and sector

The estimates in Table 5 and Table 6 represent the average estimates for the whole sample. The results,

however, mask significantly heterogeneity across regions and sectors. As a result, we further decompose the estimates in order to identify this heterogeneity. Instead of estimating the model by sub-samples we take advantage of the full sample and the structural equation estimates and estimate interactive coefficients for the innovation variables by full-maximum likelihood following Baum et al. (2015). One advantage of estimating the full model is that it allows for a straightforward test of the equality of the coefficients.

Table 7 reports the innovation coefficients only by income levels for both process innovation and for automation. Product innovation is positive and statistically significant in all regions, and we cannot reject the equality of the coefficients. On the other hand, process innovation is not statistically significant in the main or in the technological imitation specifications. The results for automation are again strong for all income regions, larger than radical product innovation and much larger than product imitation. We do not observe, however, large differences in returns across income regions.

Table 8 reports the same specifications but by geographical region. The results mimic what is observed for Table 7 but now there are no positive returns to innovation for product imitation. Also, we do not find statistically significant differences in returns across geographical regions. This contrasts with previous estimates in Cirera et al. (2015) where the returns to innovation were found to be lower in Africa and in low income countries using a dummy variable decomposition. This is likely to be the result of two main elements. First, imitation and radical innovation need to be modeled differently in the structural model. Second, the large incidence of imitation in the Africa region is likely to explain lower returns when decomposing the innovation dummy on the additional returns due to radical innovation.

Finally, Table 9 reports the estimates of innovation on productivity for different sectors, manufacturing vs services; and in the case of manufacturing, by technological intensity. Interestingly when we decompose the returns by type of technology intensity sector, we observe similar returns in high-tech sectors and low-tech sectors or in services. Again, imitation is not statistically significant, and most of the product innovation returns are related to more radical forms of innovation, also in services.

5.3.3 Sensitivity analysis

In order to test the effect of functional form in the results we also estimate the model for R&D intensity (US dollars per worker); although, this is not our preferred specification given the likely measurement error related to inputting R&D expenditure. We use a linear model for R&D since we have significant convergence problems when using a censored Tobit model. The results are qualitatively similar, and we

find positive and statistically significant high returns to innovation, especially associated to more radical forms of innovation and automation.

Also, a second test is to identify how sensitive are the results to the productivity proxy. This second set of estimates uses value added per worker instead of sales per worker as a proxy for productivity in the last stage. Although we lose a significant number of observations, the results are also qualitatively similar, and again the returns to innovation appear to be significantly high in developing countries; although mainly associated to more radical innovations.

6 Conclusion

This paper has sought to address two main gaps in the empirical literature on innovation; improving our understanding of the relationship between innovation and productivity in developing countries, and the role of different types of innovation - imitation and more radical forms of innovation. The results suggest that the returns to innovation are positive and high, but much lower (around 6 times) for imitation; which is the form of innovation more prevalent in countries furthest away from the technological frontier, especially in Sun-Saharan Africa. Indeed, when decomposing the estimates by region the returns to imitation are not statistically significant. These lower or zero returns to imitation could be the result of some measurement problems related to the definition of what constitutes an innovation and the distinction between innovation as improvement vs change; the low quality and incremental nature of some of these innovations and the lack of adequate complementary factors. In addition, we find that R&D is an important input for innovation in more radical forms and not so much for imitation.

Another important finding of the paper is the fact that in line with a large number of empirical papers in OECD countries, we find that the returns to product innovation are significantly larger than for process innovation. However, we show that when more restricted definitions of process innovation are taken including only the introduction of automation processes, the returns to process innovation increase significantly, in some cases above the returns to product innovation. Finally, we do not find statistically significant differences in the returns to innovations across regions and sectors. The findings above appear to hold when considering different sample decompositions. The returns to product innovation, radical product innovation and automation are high, while the returns to product imitation and process innovations are statistically not significant.

Overall, the results have important policy implications. The lack of returns to imitation in some regions may likely be driven by a lack of investments in some innovation activities, and, therefore,

suggest that policy needs to have a greater focus on some of the skills and complementary firm capabilities required to innovate beyond R&D, as well as some of the potential policy distortions that may be bringing down the returns to imitation.

The paper also has important implications for future research. First, we have identified significant measurement problems in terms of what constitutes an innovation, and more efforts to improve current frameworks of measuring firm-level innovation under the Oslo manual are needed. Second, more research is needed in order to understand what the critical competencies for innovation and productivity growth are, especially in poorer countries, such as managerial practices and organizational capabilities. Finally, more granularity in the analysis is needed in relation to better measure process innovation and the extent of efficiency increasing processes such as automation, and the degree of novelty in product innovation.

Table 1: List of countries

Country	Region	Income	Observations
Congo,Dem.Rep.	Africa	Low Income	529
Ghana	Africa	Low Income	720
Kenya	Africa	Low Income	713
Namibia	Africa	Higher Middle Income	379
Nigeria	Africa	Low Income	892
SouthSudan	Africa	Low Income	543
Sudan	Africa	Low Income	412
Tanzania	Africa	Low Income	723
Uganda	Africa	Low Income	640
Zambia	Africa	Low Income	720
Malawi	Africa	Low Income	250
Bangladesh	South Asia	Low Income	1442
India	South Asia	Lower Middle Income	3492
Nepal	South Asia	Low Income	482
Pakistan	South Asia	Low Income	696
Albania	Eastern Europe	Lower Middle Income	360
Armenia	Eastern Europe	Lower Middle Income	360
Azerbaijan	Eastern Europe	Lower Middle Income	390
Belarus	Eastern Europe	Higher Middle Income	360
Bosnia and Herzegovina	Eastern Europe	Lower Middle Income	360
Bulgaria	Eastern Europe	Higher Middle Income	293
Croatia	Eastern Europe	Higher Middle Income	360
Czech Republic	Eastern Europe	High Income	254
Estonia	Eastern Europe	Higher Middle Income	273
Georgia	Eastern Europe	Lower Middle Income	360
Hungary	Eastern Europe	Higher Middle Income	310
Kazakhstan	Eastern Europe	Higher Middle Income	600
Kosovo	Eastern Europe	Lower Middle Income	202
Kyrgyzstan	Eastern Europe	Low Income	270
Latvia	Eastern Europe	Higher Middle Income	336
Lithuania	Eastern Europe	Higher Middle Income	270
Macedonia	Eastern Europe	Lower Middle Income	360
Moldova	Eastern Europe	Low Income	360
Mongolia	Eastern Europe	Lower Middle Income	360
Montenegro	Eastern Europe	Higher Middle Income	173
Poland	Eastern Europe	Higher Middle Income	542
Romania	Eastern Europe	Higher Middle Income	540
Russia	Eastern Europe	Higher Middle Income	4220
Serbia	Eastern Europe	Higher Middle Income	360
Slovakia	Eastern Europe	High Income	285
Slovenia	Eastern Europe	High Income	359
Tajikistan	Eastern Europe	Low Income	369
Turkey	Eastern Europe	Higher Middle Income	1344
Ukraine	Eastern Europe	Lower Middle Income	1002
Uzbekistan	Eastern Europe	Low Income	390
Egypt	MENA	Lower Middle Income	2897
Israel	MENA	High Income	483
Jordan	MENA	Lower Middle Income	573
Lebanon	MENA	Higher Middle Income	561
Morocco	MENA	Lower Middle Income	407
Tunisia	MENA	Lower Middle Income	592
West Bank	MENA	Lower Middle Income	434
Yemen	MENA	Low Income	353

Note: Income classification based on GNI per capita and World Bank Atlas method. Source: World Bank Enterprise Surveys with Innovation modules.

Table 2: Sample overview

	Count	Share
Size		
small($¡$20)	17,594	50.3
medium(20 to 99)	11,787	33.7
large(100 and over)	5,574	16
Industry		
Food	3,127	9.2
Textiles	1,177	3.5
Garments	2,030	5.9
Wood, Paper	852	2.5
Publishing, Printing	912	2.7
Chemicals	1,244	3.6
Plastics	998	2.9
Non metallic mineral products	1,184	3.5
Basic metals, products	2,106	6.2
Machinery	981	2.9
Electronics	1,219	3.6
Furniture	1,175	3.4
Manufacturing total	17,005	49.8
Construction	1,986	5.8
Motor vehicle services	1,073	3.1
Transportation	1,628	4.8
Wholesale	3,391	9.9
Retail	6,349	18.6
Hotels, Restaurants	2,112	3.5
IT, Professional	611	1.8
Services total	17,150	50.2

Source: World Bank Enterprise Surveys with Innovation modules.

Table 3: Variable Description

Variables	Description
Knowledge intensity	
R&D dummy	Dummy with value 1 if firm invests in Intramural or extramural R&D
R&D intensity	(RD+/L) Intramural and extramural R&D expenditure per worker
Innovation outputs	
Product/process innovation	Dummy with value 1 if any new or significantly improved product, service or process introduced by this establishment in last three
years Product innovation	Dummy with value 1 if any new or significantly improved product or service introduced by this establishment in last three years.
Process innovation	Dummy with value 1 if any new or significantly improved processes introduced by this establishment in last three years.
Productivity	
Sales per worker (sales/L)	Logarithm of sales per worker
Firms Market Condition and Access to Finance	
Working capital	Share of working capital financed by internal funds. This is a proxy to measure the degree of external financial constrain for the
firm Duopoly/monopoly	Whether firms face one or two main competitors in the market
External Market (Exporter and Importer)	Export and Import dummies
Informal Sector Competition	How much practices of informal sector is an obstacle. Index 0 not an obstacle to 4 severe obstacle.
Demand (Increasing) pull effect	Dummy indicating whether firms demand has increased by evaluating revenue or employment
growth ***Technology Push Factors***	
License Foreign Technology	Dummy whether a firm use technology licensed from a foreign owned company
New Capital in Previous Year	Dummy whether a firm has purchased any new fixed asset in the last fiscal year
Firm characteristics	
Log (K/L)	Capital intensity defined as the log of the ratio of capital to labor in the firm
Educational Obstacles	Dummy created for those which inadequately educated workforce presents major or severe
obstacles Age	Log of firm's age
Log(L) - Size	Log of employment,number of workers
Agglomeration	
Spillovers	This is the share of other innovators in the same region and sector
Business city	Dummy with value 1 if location of the establishment is in the main business city
Other Controls	
Sector Dummy	Dummy based on SIC classification
Country Dummy	Dummy for each country

Table 4: Benchmarking innovation rates

	Product Innovation			Process Innovation			Organizational Innovation
	All	Imitation	Radical	All	Imitation	Radical	All
Africa	**26%**	**22%**	**4%**	**31%**	**26%**	**4%**	**43%**
Congo, Dem. Rep.	19%	18%	2%	23%	20%	2%	40%
Ghana	16%	13%	3%	24%	22%	3%	31%
Kenya	25%	20%	6%	26%	22%	5%	37%
Namibia	37%	25%	13%	53%	42%	11%	65%
Nigeria	13%	12%	2%	29%	25%	4%	47%
South Sudan	49%	40%	9%	35%	28%	7%	46%
Sudan	2%	2%	1%	4%	4%	1%	20%
Tanzania	21%	20%	0%	17%	16%	1%	26%
Uganda	34%	32%	3%	34%	32%	1%	71%
Zambia	33%	30%	3%	51%	47%	4%	36%
Malawi	33%	30%	3%	41%	35%	7%	51%
Asia	**29%**	**25%**	**4%**	**35%**	**31%**	**4%**	**36%**
Bangladesh	44%	37%	7%	61%	55%	7%	37%
India	53%	47%	6%	61%	54%	7%	55%
Nepal	12%	12%	0%	10%	9%	1%	37%
Pakistan	8%	6%	2%	8%	6%	2%	15%
ECA	**19%**	**10%**	**10%**	**13%**	**8%**	**5%**	**53%**
Albania	8%	5%	3%	2%	1%	1%	34%
Armenia	10%	7%	3%	3%	3%	0%	85%
Azerbaijan	0%	0%	0%	2%	2%	0%	77%
Belarus	25%	12%	13%	26%	21%	5%	53%
Bosnia and Herzegovina	35%	13%	22%	23%	14%	9%	72%
Bulgaria	21%	13%	7%	14%	11%	3%	54%
Croatia	32%	17%	15%	11%	7%	3%	47%
Czech Republic	30%	15%	15%	18%	12%	6%	59%
Estonia	13%	8%	5%	14%	12%	2%	45%
Georgia	5%	3%	2%	7%	5%	2%	49%
Hungary	13%	7%	6%	12%	4%	7%	31%
Kazakhstan	13%	6%	7%	10%	5%	5%	64%
Kosovo	49%	19%	30%	27%	8%	20%	83%
Kyrgyz Republic	34%	7%	27%	23%	7%	16%	77%
Latvia	13%	9%	4%	7%	7%	1%	42%
Lithuania	15%	4%	12%	11%	4%	7%	54%
Macedonia, FYR	20%	8%	12%	14%	10%	4%	48%
Moldova	27%	11%	15%	14%	5%	10%	73%
Mongolia	14%	5%	9%	19%	10%	9%	56%
Montenegro	21%	16%	4%	17%	10%	7%	31%
Poland	18%	10%	9%	11%	7%	4%	49%
Romania	32%	24%	8%	24%	19%	4%	49%
Russian Federation	21%	12%	9%	17%	10%	6%	68%
Serbia	22%	14%	8%	13%	9%	4%	44%
Slovak Republic	11%	7%	4%	12%	11%	1%	34%
Slovenia	13%	6%	7%	8%	5%	3%	51%
Tajikistan	45%	17%	28%	6%	3%	3%	53%
Turkey	5%	4%	1%	6%	5%	1%	63%
Ukraine	7%	6%	1%	7%	6%	1%	45%
Uzbekistan	4%	2%	2%	0%	0%	0%	10%
MENA	**16%**	**6%**	**10%**	**12%**	**7%**	**9%**	**58%**
Egypt	9%	5%	4%	8%	5%	3%	40%
Israel	14%	3%	11%	3%	3%	2%	66%
Jordan	8%	5%	3%	6%	4%	4%	24%
Lebanon	28%	9%	19%	17%	9%	13%	59%
Morocco	13%	4%	8%	14%	6%	17%	59%
Tunisia	22%	6%	15%	23%	11%	13%	69%
West Bank	11%	6%	5%	10%	7%	6%	59%
Yemen	23%	11%	12%	11%	9%	13%	87%

Source: Authors elaboration from Enterprise Surveys innovation modules

Table 5: Returns of Innovation on Productivity, Full Model

	Product & process		Product & process (novelty)			Technological	Technological (novelty)	
	(1a)	(1b)	(2a)	(2b)	(2c)	(3)	(4a)	(4b)
Sales per worker								
product/technological	1.143***					0.852***		
	(0.06)					(0.07)		
process	0.214**		0.488***					
	(0.09)		(0.09)					
K/L Ratio	0.172***		0.172***			0.173***	0.174***	
	(0.01)		(0.01)			(0.01)	(0.01)	
Firms Size - in log	0.092***		0.091***			0.098***	0.106***	
	(0.01)		(0.01)			(0.01)	(0.01)	
imitation			0.394***				0.155**	
			(0.10)				(0.07)	
radical			0.995***				0.740***	
			(0.09)				(0.08)	
Constant	11.895***		11.962***			11.821***	12.058***	
	(0.13)		(0.13)			(0.13)	(0.13)	
innovation incidence	**product**	**process**	**imitation**	**radical**	**process**	**technology**	**tech_imitation**	**tech_radical**
Invests in R&D	0.657***	0.626***	0.182*	0.342***	0.463***	0.836***	0.015	0.372***
	(0.10)	(0.11)	(0.11)	(0.10)	(0.13)	(0.09)	(0.09)	(0.08)
Firms Size - in log	0.016*	0.058***	0.005	0.053***	0.062***	0.043***	0.016*	0.068***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Log Firm Age	-0.008	0.013	-0.037***	0.035**	0.014	0.003	-0.024**	0.028**
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Education as Obstacle	0.102***	0.166***	0.075***	0.062**	0.164***	0.139***	0.084***	0.077***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)
Importer	0.175***	0.119***	0.018	0.206***	0.143***	0.193***	-0.017	0.206***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Firm Exports	0.169***	0.100***	0.082***	0.214***	0.120***	0.148***	0.037	0.186***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)
Demand Pull Effect	0.165***	0.099***	0.126***	0.086***	0.112***	0.145***	0.096***	0.073***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Duopoly, monopoly	-0.015	-0.048	-0.014	-0.001	-0.048*	-0.054**	-0.077***	0.015
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Working K	-0.001***	-0.002***	-0.001***	-0.000	-0.002***	-0.002***	-0.001***	-0.001**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Business city	0.086***	-0.036	-0.024	0.169***	-0.023	0.053***	-0.069***	0.166***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)
spillover	1.949***	1.858***	1.055***	1.566***	1.922***	2.343***	1.312***	1.719***
	(0.31)	(0.33)	(0.36)	(0.41)	(0.33)	(0.29)	(0.32)	(0.38)
R&D incidence								
Firms Size - in log	0.148***		0.154***			0.149***	0.156***	
	(0.01)		(0.01)			(0.01)	(0.01)	
Log Firm Age	0.019		0.019			0.018	0.019	
	(0.01)		(0.01)			(0.01)	(0.01)	
importer	0.134***		0.140***			0.130***	0.136***	
	(0.03)		(0.03)			(0.03)	(0.03)	
Firm Exports	0.283***		0.285***			0.281***	0.284***	
	(0.03)		(0.03)			(0.03)	(0.03)	
New Capital in Previous Year	0.308***		0.274***			0.309***	0.259***	
	(0.03)		(0.03)			(0.02)	(0.02)	
Informal Sector as Obstacle	0.022		0.019			0.025	0.019	
	(0.03)		(0.03)			(0.03)	(0.03)	
license foreign	0.298***		0.268***			0.296***	0.263***	
	(0.03)		(0.03)			(0.03)	(0.03)	
working K	-0.002***		-0.002***			-0.001***	-0.002***	
	(0.00)		(0.00)			(0.00)	(0.00)	
duopoly monopoly	-0.032		-0.035			-0.032	-0.034	
	(0.03)		(0.03)			(0.03)	(0.03)	
N	30798		30798			30798	30798	
Pseudo log-likelihood	-75480.12		-79470.39			-64195.45	-70282.09	

Note: Coefficients and standard errors robust to heteroskedasticity and 1, 5, and 10 percent levels of significance are denoted by ***, ** and *, respectively. All estimates include country and sector fixed effects, not shown. Marginal effects shown for Innovation stage. Columns 1a,1b present results from the model where product and process innovation are introduced separately. Columns 2a,2b,2c present the model where product innovation is classified by novelty (imitation vs. radical). Column 3 presents the model where product and process innovation are combined as technological innovation. Column 4a, 4b present the model where technological innovation is classified by novelty.

Table 6: Returns of Innovation on Productivity, Full Model with Process Automation

	Product & automation		Product & automation (novelty)			Technological	Technological (novelty)	
	(1a)	(1b)	(2a)	(2b)	(2c)	(3)	(4a)	(4b)
Sales per worker								
product/technological	1.043***					1.184***		
	(0.07)					(0.05)		
process	1.261***		1.328***					
	(0.07)		(0.07)					
K/L Ratio	0.171***		0.171***			0.173***	0.172***	
	(0.01)		(0.01)			(0.01)	(0.01)	
Firms Size - in log	0.093***		0.100***			0.097***	0.103***	
	(0.01)		(0.01)			(0.01)	(0.01)	
imitation			0.330***				0.304***	
			(0.09)				(0.08)	
radical			0.905***				1.071***	
			(0.09)				(0.06)	
Constant	11.904***		12.017***			11.887***	12.042***	
	(0.13)		(0.13)			(0.13)	(0.13)	
innovation incidence	**product**	**process**	**imitation**	**radical**	**process**	**technology**	**tech imitation**	**tech radical**
Invests in R&D	0.619***	0.354***	0.265***	0.401***	0.345***	0.673***	0.010	0.349***
	(0.09)	(0.13)	(0.09)	(0.09)	(0.13)	(0.09)	(0.09)	(0.08)
Firms Size - in log	0.019**	0.003	0.001	0.053***	0.002	0.015*	-0.003	0.043***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Log Firm Age	-0.011	0.028	-0.038***	0.030*	0.025	-0.006	-0.043***	0.027**
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Education as Obstacle	0.105***	0.050	0.074***	0.070**	0.056	0.100***	0.061**	0.066***
	(0.02)	(0.04)	(0.03)	(0.03)	(0.04)	(0.02)	(0.03)	(0.03)
importer	0.169***	0.120***	0.003	0.202***	0.128***	0.162***	0.006	0.126***
	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
Firm Exports	0.169***	0.119***	0.070***	0.211***	0.131***	0.181***	0.080***	0.202***
	(0.02)	(0.04)	(0.03)	(0.03)	(0.04)	(0.02)	(0.03)	(0.03)
Demand Pull Effect	0.161***	0.144***	0.123***	0.081***	0.157***	0.178***	0.120***	0.120***
	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
duopoly,monopoly	-0.014	0.021	-0.017	0.003	0.020	-0.015	-0.026	0.011
	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)
working K	-0.001***	-0.002***	-0.001***	-0.000	-0.002***	-0.001***	-0.000	-0.001***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
business	0.080***	-0.054	-0.031	0.171***	-0.050	0.065***	-0.019	0.096***
	(0.02)	(0.04)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
spillover	1.840***	2.664***	0.993***	1.461***	2.775***	2.317***	0.796**	2.071***
	(0.31)	(0.62)	(0.36)	(0.42)	(0.61)	(0.30)	(0.37)	(0.37)
Constant	-1.187***	-1.568***	-0.931***	-2.612***	-1.560***	-1.017***	-0.965***	-1.794***
	(0.09)	(0.14)	(0.09)	(0.16)	(0.14)	(0.09)	(0.09)	(0.11)
R&D incidence								
Firms Size - in log	0.149***		0.152***			0.149***	0.154***	
	(0.01)		(0.01)			(0.01)	(0.01)	
Log Firm Age	0.019		0.018			0.019	0.019	
	(0.01)		(0.01)			(0.01)	(0.01)	
importer	0.134***		0.132***			0.133***	0.134***	
	(0.03)		(0.03)			(0.03)	(0.03)	
Firm Exports	0.282***		0.282***			0.282***	0.284***	
	(0.03)		(0.03)			(0.03)	(0.03)	
New Capital in Previous Year	0.306***		0.292***			0.306***	0.275***	
	(0.02)		(0.02)			(0.02)	(0.02)	
Informal Sector as Obstacle	0.021		0.023			0.022	0.024	
	(0.03)		(0.03)			(0.03)	(0.03)	
license foreign	0.294***		0.278***			0.297***	0.270***	
	(0.03)		(0.03)			(0.03)	(0.03)	
working K	-0.001***		-0.002***			-0.001***	-0.002***	
	(0.00)		(0.00)			(0.00)	(0.00)	
duopoly monopoly	-0.032		-0.033			-0.033	-0.032	
	(0.03)		(0.03)			(0.03)	(0.03)	
N	30798		30798			30798	30798	
Pseudo log-likelihood	-67515.00		-71484.82			-63248.83	-68415.04	

Note: Coefficients and standard errors robust to heteroskedasticity and 1, 5, and 10 percent levels of significance are denoted by ***, ** and *, respectively. All estimates include country and sector fixed effects, not shown. Marginal effects shown for Innovation stage. Columns follow similar approach to Table 5. Process innovation restricted to automation.

Table 7: Returns of Innovation on Productivity, by GDP

Probit model	One sample interaction (prod and proc)						One sample interaction (automation)					
	low income	middle income	high income	low income	middle income	high income	low income	middle income	high income	low income	middle income	high income
product	1.189***	1.109***	1.136***				1.052***	0.987***	1.030***			
	-0.07	-0.06	-0.07				(0.08)	(0.07)	(0.07)			
process	0.134	0.158*	0.11	0.414***	0.446***	0.408***	1.198***	1.273***	1.218***	1.287***	1.339***	1.301***
	-0.09	-0.09	-0.09	-0.09	-0.1	-0.1	(0.09)	(0.08)	(0.10)	(0.09)	(0.07)	(0.09)
imitation				0.410***	0.410***	0.384***				0.299***	0.311***	0.295***
				-0.1	-0.1	-0.11				(0.09)	(0.09)	(0.10)
radical				1.119***	0.924***	0.928***				1.010***	0.802***	0.849***
				-0.11	-0.09	-0.09				(0.11)	(0.10)	(0.09)
technological	0.886***	0.850***	0.814***				1.240***	1.159***	1.138***			
	(0.07)	(0.07)	(0.07)				(0.06)	(0.06)	(0.06)			
tech imitation				0.144*	0.141*	0.029				0.355***	0.349***	0.224**
				(0.07)	(0.08)	(0.08)				(0.09)	(0.08)	(0.09)
tech radical				0.806***	0.631***	0.679***				1.116***	1.049***	0.978***
				(0.09)	(0.08)	(0.08)				(0.07)	(0.06)	(0.07)

Note: Coefficients and standard errors robust to heteroskedasticity and 1, 5, and 10 percent levels of significance are denoted by ***, ** and *, respectively. All estimates include country and sector fixed effects, not shown.

Table 8: Returns of Innovation on Productivity, by Region

Probit model	One sample interaction (product, process)								One sample interaction (product, automation)							
	Africa	Asia	ECA	MENA	Africa	Asia	ECA	MENA	Africa	Asia	ECA	MENA	Africa	Asia	ECA	MENA
product	1.094***	1.007***	1.099***	1.018***					0.865***	0.758***	0.902***	0.813***				
	(0.08)	(0.07)	(0.07)	(0.08)					(0.09)	(0.09)	(0.08)	(0.09)				
process	-0.057	0.090	-0.018	-0.065	0.167*	0.316***	0.216**	0.196*	1.173***	1.303***	1.214***	1.601***	1.218***	1.333***	1.266***	1.662***
	(0.09)	(0.10)	(0.10)	(0.11)	(0.10)	(0.11)	(0.11)	(0.11)	(0.12)	(0.07)	(0.09)	(0.16)	(0.12)	(0.07)	(0.08)	(0.16)
imitation					0.204	0.220*	0.202	0.195					0.199*	0.192*	0.219*	0.203
					(0.13)	(0.13)	(0.14)	(0.15)					(0.11)	(0.10)	(0.11)	(0.13)
radical					1.195***	0.983***	1.060***	0.870***					1.052***	0.738***	0.908***	0.729***
					(0.13)	(0.10)	(0.09)	(0.10)					(0.13)	(0.11)	(0.09)	(0.10)
technological	0.600***	0.719***	0.652***	0.617***					0.600***	0.719***	0.652***	0.617***				
	(0.08)	(0.09)	(0.07)	(0.08)					(0.08)	(0.09)	(0.07)	(0.08)				
tech imitation					0.015	0.173*	-0.038	-0.010					0.015	0.173*	-0.038	-0.010
					(0.08)	(0.09)	(0.09)	(0.10)					(0.08)	(0.09)	(0.09)	(0.10)
tech radical					0.666***	0.677***	0.720***	0.575***					0.666***	0.677***	0.720***	0.575***
					(0.10)	(0.09)	(0.08)	(0.09)					(0.10)	(0.09)	(0.08)	(0.09)

Note: Coefficients and standard errors robust to heteroskedasticity and 1, 5, and 10 percent levels of significance are denoted by ***, ** and *, respectively. All estimates include country and sector fixed effects, not shown.

Table 9: Returns of Innovation on Productivity, by Sector

Probit model	One sample interaction - product and process								One sample interaction - product and automation							
	High	**Medium**	**Low**	**Services**	**High**	**Medium**	**Low**	**Services**	**High**	**Medium**	**Low**	**Services**	**High**	**Medium**	**Low**	**Services**
product	1.107***	1.175***	1.139***	1.252***					0.997***	1.069***	1.038***	1.169***				
	(0.07)	(0.07)	(0.07)	(0.06)					(0.08)	(0.08)	(0.07)	(0.07)				
process	0.210**	0.198**	0.228***	0.152*	0.505***	0.519***	0.534***	0.478***	1.219***	1.238***	1.229***	1.438***	1.268***	1.319***	1.311***	1.543***
	(0.10)	(0.09)	(0.09)	(0.09)	(0.10)	(0.10)	(0.09)	(0.09)	(0.09)	(0.09)	(0.08)	(0.12)	(0.09)	(0.09)	(0.08)	(0.12)
imitation					0.157	0.234**	0.129	0.163					0.153	0.243**	0.139	0.193**
					(0.11)	(0.12)	(0.11)	(0.11)					(0.10)	(0.10)	(0.10)	(0.10)
radical					1.143***	1.012***	1.100***	1.266***					1.009***	0.877***	0.970***	1.166***
					(0.09)	(0.10)	(0.09)	(0.08)					(0.10)	(0.11)	(0.10)	(0.09)
Technological	0.930***	0.905***	0.891***	0.912***					1.160***	1.180***	1.192***	1.272***				
	(0.08)	(0.08)	(0.07)	(0.07)					(0.07)	(0.07)	(0.06)	(0.06)				
tech imitation					0.214**	0.180**	0.137*	0.101								
					(0.09)	(0.09)	(0.08)	(0.08)								
tech radical					0.846***	0.718***	0.769***	0.794***								
					(0.10)	(0.10)	(0.09)	(0.09)								

Note: Coefficients and standard errors robust to heteroskedasticity and 1, 5, and 10 percent levels of significance are denoted by ***, ** and *, respectively. All estimates include country and sector fixed effects, not shown. Columns label High, medium, low refer to the technological intensity within the manufacturing sector.

Figure 1: Share of Innovators Using R&D, by country

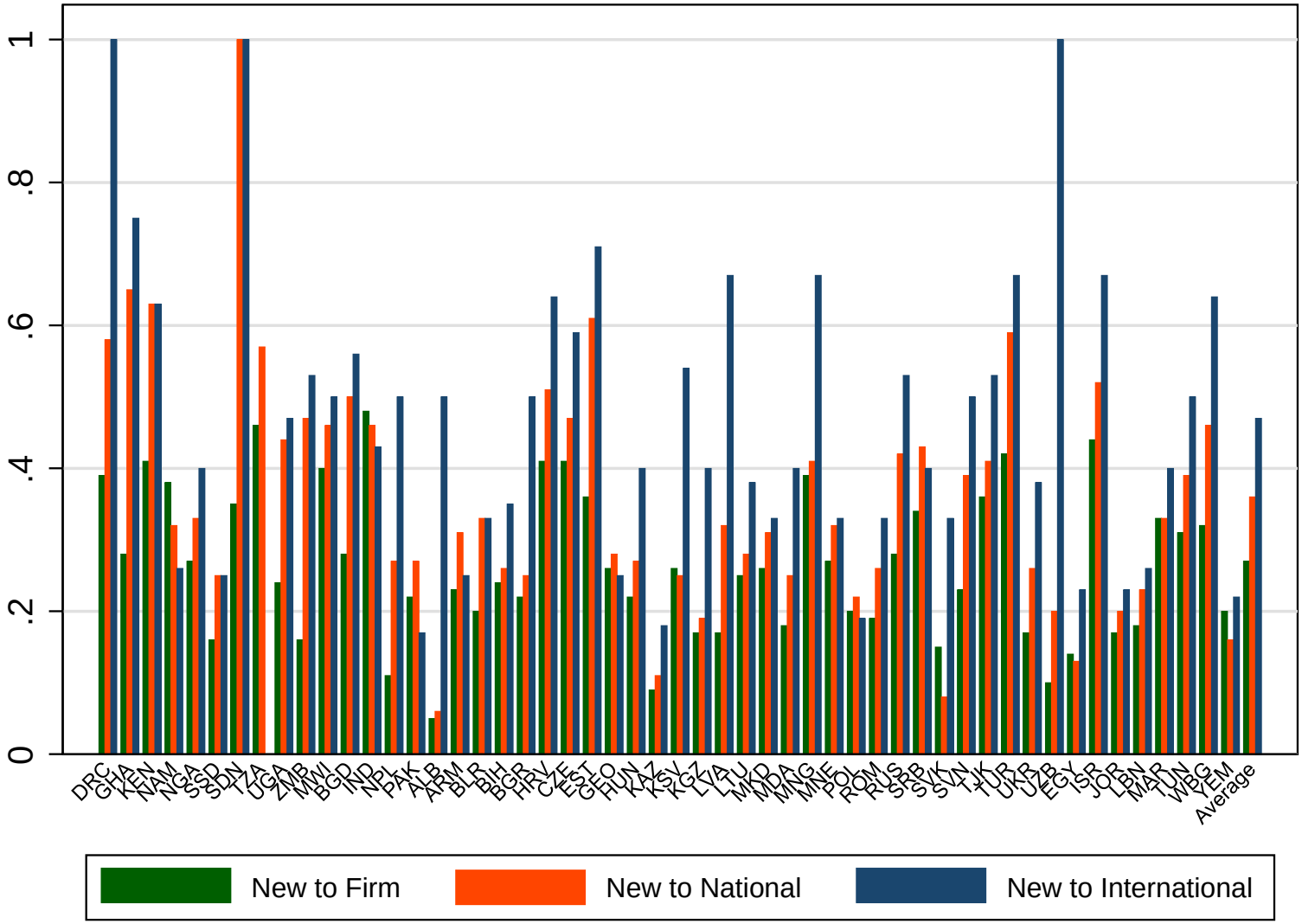


Figure 2: Average R&D expenditures by GDP

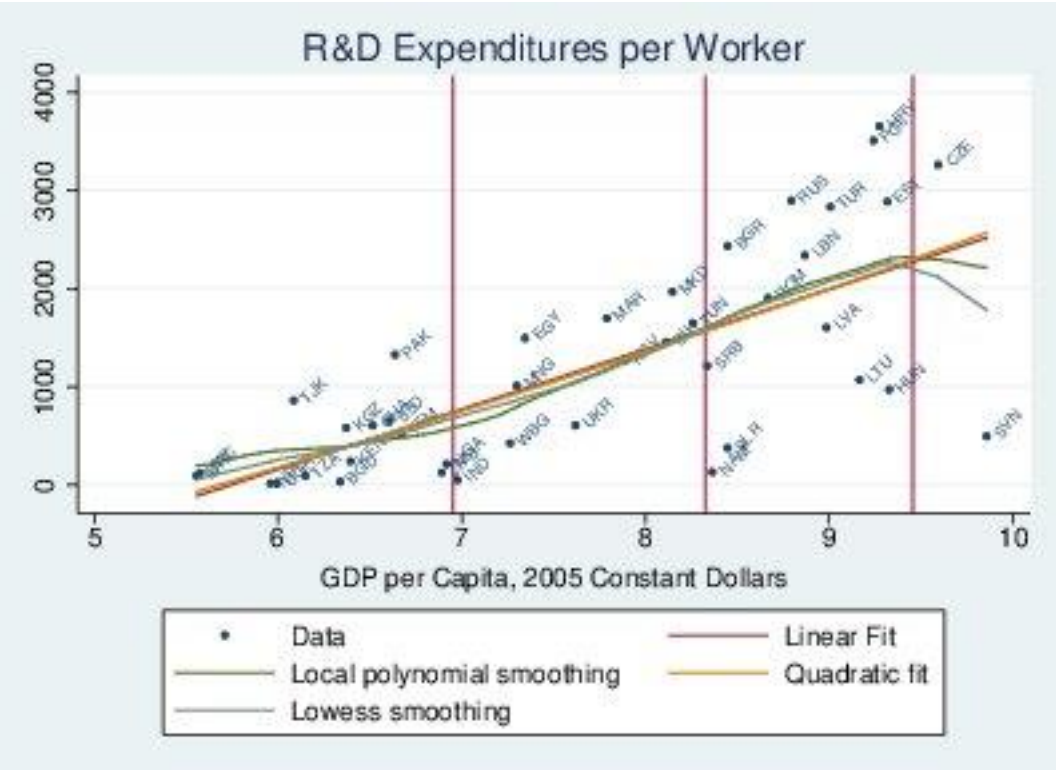
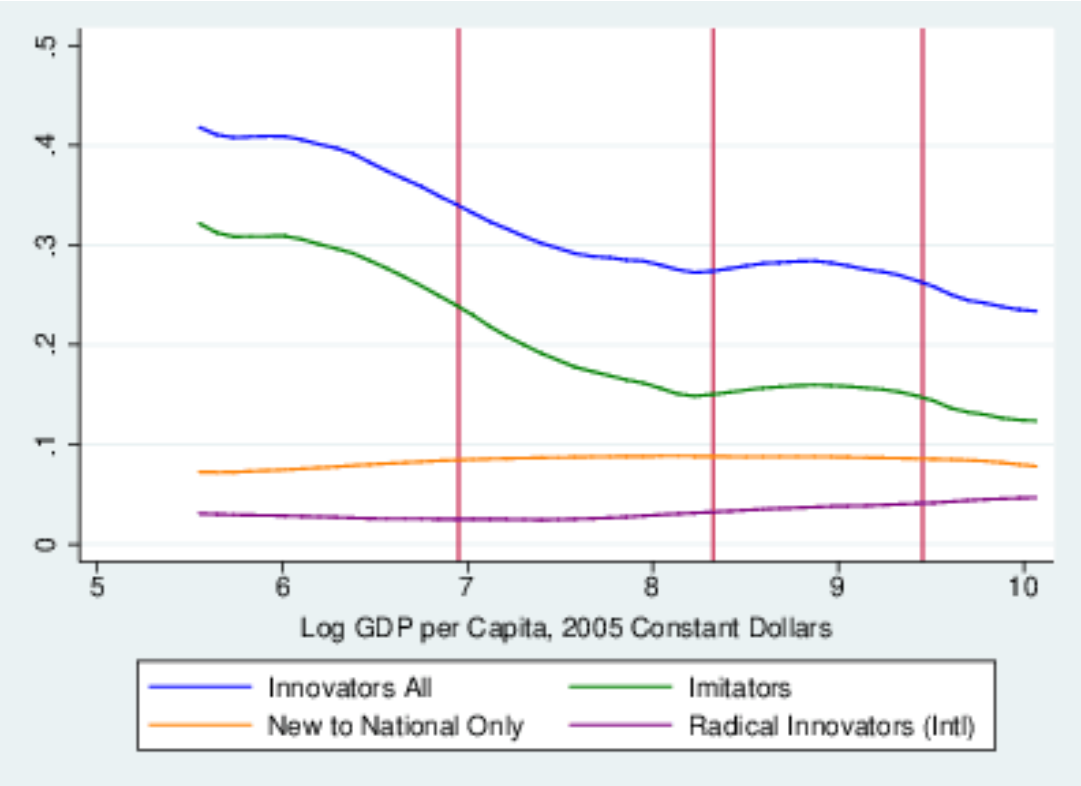


Figure 3: Innovation rates by GDP



A Definitions of Innovation

A challenge of measuring innovation outcomes is the subjective nature of many of the questions used in the surveys. The Oslo manual, which is the main reference for these type of surveys, defines innovation as "..the implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organizational method in business practices, workplace organization or external relations." The Enterprise Survey (ES) uses this definition to identify innovations by directly asking firm managers and owners whether they have implemented "new" or "significant" changes or improvements in the last three years. However, this approach raises additional complexity for the researcher as "significant" is highly subjective and self-reported, and implies some confusion about identifying innovation. As a result, significant cleaning and re-classification of the data based on the written definitions of product and process innovations was performed to improve the quality of the data; however, the issue of measurement remains an open area for improvement. The definitions of innovation in the survey questionnaire were based on the same interpretation as in the Community Innovation Survey and Oslo Manual, outlined below.

- Product Innovation A product innovation is the introduction of a good or service that is new or significantly improved with respect to its characteristics or intended uses. This includes significant improvements in technical specifications, components and materials, incorporated software, user friendliness or other functional characteristics. Product innovations can utilize new knowledge or technologies, or can be based on new uses or combinations of existing knowledge or technologies.

- Process Innovation Process innovation is the implementation of a new or significantly improved production or delivery method. This includes significant changes in techniques, equipment and/or software. Process innovations can be intended to decrease unit costs of production or delivery, to increase quality, or to produce or deliver new or significantly improved products. In the questionnaire, process innovation is defined as one of the following:

    1. Automate manual processes, partially or fully
    2. Adapt a technology or method previously used by this establishment?
    3. Introduce a new technology or method

4. Use a more efficient technology or method already used by this establishment

- Organizational Innovation

  1. Outsourcing tasks & changes in relations
  2. Changes in management structures, integrating departments/units
  3. Acquiring management systems for information & knowledge

B Re-classification of innovation

We re-classified product and process innovations based on their descriptions provided in the questionnaire in order to correct for miss-classification. In some cases, there was not enough information to validate an innovation, and in other cases innovations were misidentified between product and process innovations or process and marketing innovations. Below are some examples of how the re-classification was implemented. Figure A1 reports self-reported and cleaned innovation rates for each country in our sample.

1. Delivery

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

2. Distinction between introducing new types of product as product innovation or marketing

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

3. Other

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

Table A.1: Returns of Innovation on Productivity, Full Model and Uncleaned data

	Product & process		Technological		Product & automation		Technological (prod&automation)	
	Baseline	**Novelty**	**Baseline**	**Novelty**	**Baseline**	**Novelty**	**Baseline**	**Novelty**
Sales per worker								
product/technological	0.698***		0.445***		0.877***		0.639***	
	(0.09)		(0.06)		(0.07)		(0.07)	
process	0.400***	0.561***			0.440***	0.498***		
	(0.09)	(0.08)			(0.08)	(0.08)		
K/L Ratio	0.172***	0.172***	0.173***	0.174***	0.172***	0.173***	0.173***	0.173***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Firms Size - in log	0.088***	0.092***	0.112***	0.117***	0.093***	0.105***	0.107***	0.115***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
imitation		0.095		0.110*		0.191***		0.146**
		(0.08)		(0.06)		(0.07)		(0.07)
radical		0.531***		0.421***		0.679***		0.431***
		(0.09)		(0.08)		(0.08)		(0.06)
Constant	11.768***	11.908***	11.908***	12.045***	11.763***	11.958***	11.864***	11.994***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
N	30798	30798	30798	30798	30798	30798	30798	30798
Pseudo log-likelihood	-77569.21	-83127.23	-63395.63	-70287.17	-69946.78	-75418.93	-63650.56	-70919.41

Note: Coefficients and standard errors robust to heteroskedasticity and 1, 5, and 10 percent levels of significance are denoted by ***, ** and *, respectively. All estimates include country and sector fixed effects, not shown. Marginal effects shown for Innovation stage.

Figure A.1: Innovation rates, self-reported and cleaned

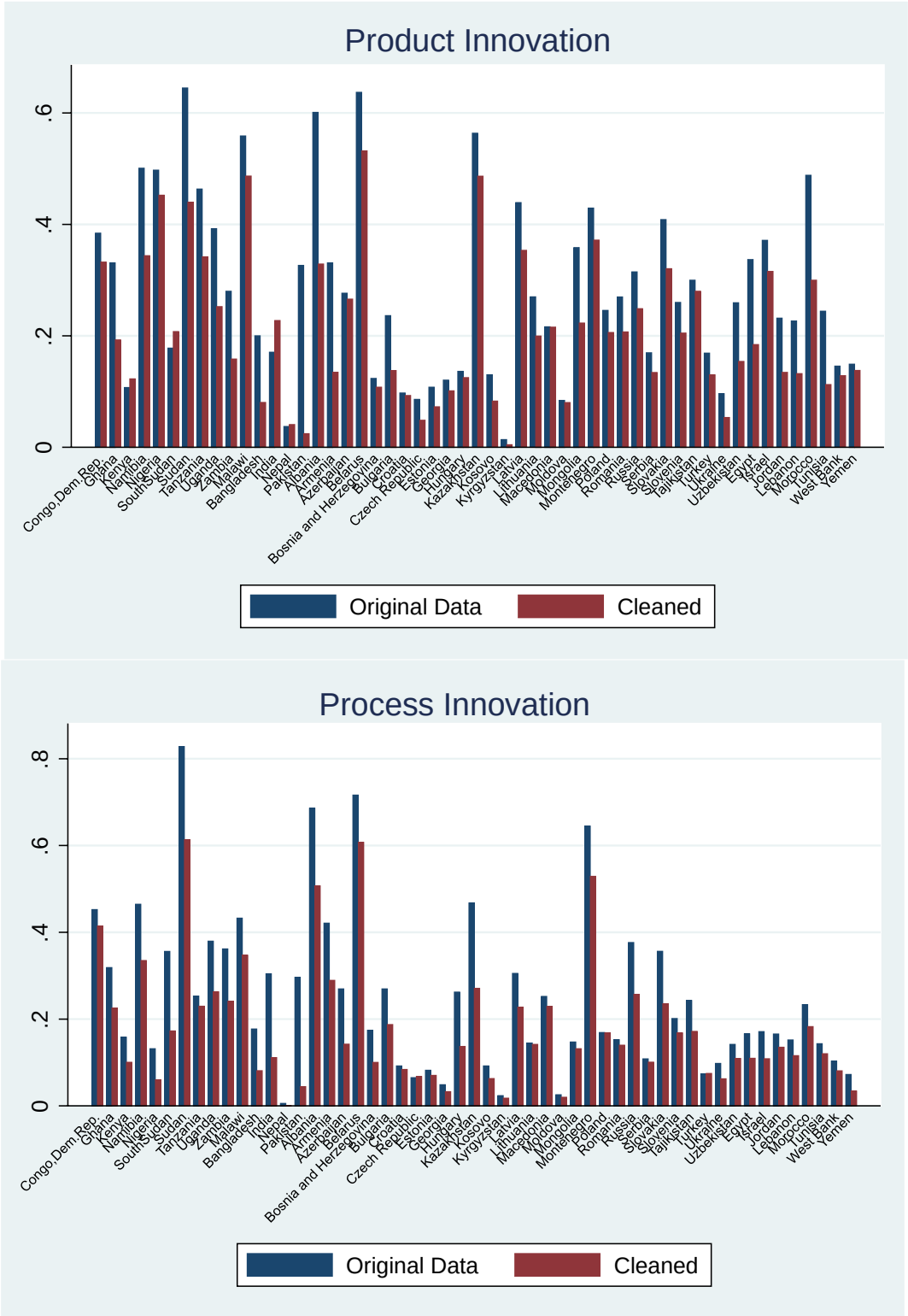
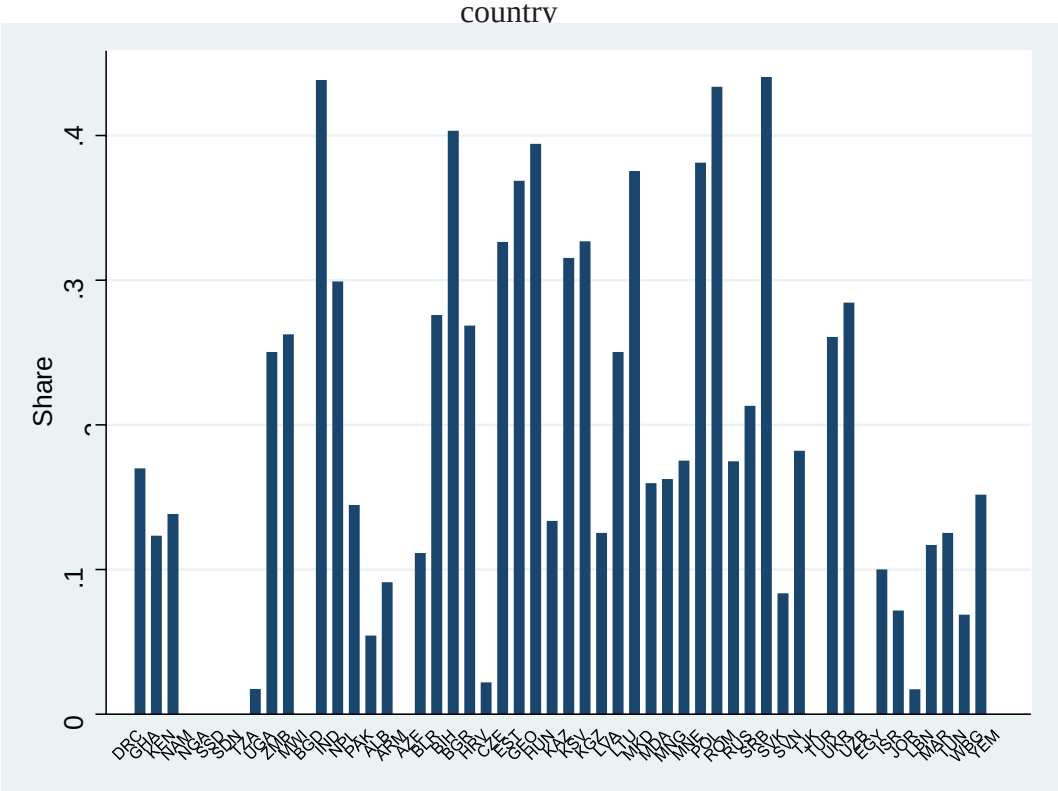


Figure A.2: Automation rates as share of Process Innovation, by country



The Effects of Innovation on Employment in Developing Countries. Evidence from Enterprise Surveys[1,2]

Xavier Cirera
The World Bank Group, Washington, D.C.
Corresponding author: xcirera@worldbank.org

Leonard Sabetti

Université d'Auvergne, Clermont Ferrand (France)

leonard.sabetti@etu.udamail.fr

[1] We are grateful for helpful comments seminar participants at the 10th IZA/World Bank conference on Technological Change and Jobs (Bonn, Germany), the 6th ZEW/MaCCI Conference on the Economics of Innovation and Patenting (Mannheim, Germany) and this volume editors. We would like to thank Jordi Jaumeandreu for extensive comments to an earlier version of this paper, Jean-Louis Combes for discussing our paper and sharing comments at a seminar presentation at CERDI, and CIIP for financial assistance. Corresponding Author: The World Bank, Firm Capabilities and innovation, Finance, Competitiveness and Innovation Global Practice, Washington, D.C. Email: xcirera@worldbank.org
[2] Published in *Industrial and Corporate Change (2019).* Please see full reference in Bibliography.

The Effects of Innovation on Employment in Developing Countries: Evidence from Enterprise Surveys

Abstract

This paper sheds light on the direct impact of technological as well as organizational innovation on firm level employment growth using a global sample of over 15,000 firms in developing countries. The main findings suggest that new sales associated with product innovation are produced, on average, with just as much or higher levels of labor intensity than old products. However, the additionality to employment decreases with productivity, proxied by income per capita. In line with other studies, process innovation does not impact the additionality of employment, but there is some evidence of automation reducing the impact of product innovation on employment.

1 Introduction

Innovation is the engine of the creative destruction process that spurs economic dynamism and transformation (Schumpeter 1942). At the macro level, theories of economic growth put innovation at the center of the growth process since Solow's (1957) seminal work, where economic growth is driven by technical change. The emergence of new growth theory emphasized the role of innovation and knowledge accumulation in the growth process and Schumpeterian creative destruction arising from a competitive R&D sector as the main engine of economic growth (Aghion and Howitt (1992); Romer, 1986).

At the firm level, Klette and Kortum (2004) show how innovation activities create rich firm-level dynamics that translate into firm growth. Unlike previous models where firm growth was largely driven by firm learning, in the model of Klette and Kortum (2004), innovation increases product quality and makes firms more competitive, which increases their revenue and size and forces existing firms producing old and obsolete versions of the product to exit the market.

While innovation has the potential to generate large productivity gains and significantly improve allocative efficiency,[3] the short-term direct impacts of innovation on employment remain an important policy question, especially in a developing country context where firms operate far from the technological frontier. While the introduction of new product lines tends to generate employment, new processes using more modern technologies or upgraded products can result in more efficient use of labor or labor replacement. Determining what the tradeoff might be (if any) between innovation and employment growth is critical for policy, especially in developing countries where the needs to absorb new entrants to the labor market in formal and higher productivity jobs are greatest.

[3] Lentz and Mortensen (2008), using Danish firm-level data, find that up to 75% of productivity growth comes from reallocation of inputs to innovating firms, of which 25% is entry and exist of firms and 50% reallocation to growing innovative firms.

This paper tries to shed some light to this tradeoff and estimates the short-run direct impact of firm-level innovation on employment in a sample of developing countries, using a novel dataset on firm innovation activities for a large set of developing countries. Our contribution to the literature is threefold. First, we expand the existing evidence on employment effects to a large number of low and middle-income countries, taking advantage of a unique firm survey implemented in 53 countries by the World Bank Enterprise Surveys Unit. Second, we attempt to disentangle the effects of process innovation from product innovation on employment when both are implemented simultaneously, as well as isolating the impact of automation from process innovation more broadly. Finally, we study how the degree of novelty of innovation – new to the firm vs new to national or international market - and the level of a country's development, as a proxy for aggregate productivity, affect the elasticity of innovation to employment. Given the nature of the data we use, our results focus on average, short-run effects at the firm-level; an important building block to arriving at aggregate or sector level effects on employment, which inform welfare considerations.

This paper is structured as follows. The next section briefly summarizes the existing evidence, mainly for OECD countries, regarding the impact of innovation on employment. Section 3 describes the data, while section 4 develops the methodological framework used in the empirical section. Section 5 describes the main empirical findings. The last section concludes.

2 The Direct Impact of Innovation and Employment

Innovation is the outcome of firms' investments in knowledge capital, managerial practices and organizational decisions. The ultimate objective of these investments is to introduce new or upgraded products or processes that positively impact firm performance by increasing productivity, sales, profits or markups. However, there is uncertainty regarding the extent to which firms are able to convert knowledge capital investments into innovation outcomes and furthermore, whether these innovation outcomes are likely to impact firm performance. This uncertainty is particularly high in developing countries, where there is a lack of key complementary factors such

as skills, managerial and organizational or technology capabilities to support innovation. The impact of innovation on employment depends on the allocation of workers complementing innovations, and the impact of these innovation efforts on firms' efficiency.

To date, most of the evidence on the impact of innovation on employment has focused on developed countries. Some of the case study literature has emphasized the role of innovation as a force for skill-biased technological change; since increases in firm efficiency can result in more efficient use of labor and changes in the relative demand for skilled labor. Few empirical studies, however, have found a direct negative impact on employment; although the bulk of studies suggest strong evidence of skill-biased technological change. From a theoretical standpoint, predicting the effects of innovation on employment can be ambiguous. While product innovation is typically aimed at increasing a firm's demand through the introduction of a new product, process innovation usually entails production enhancements that can be labor saving. The existence of these competing mechanisms makes the net effects uncertain.

A strand of the literature has analyzed the impact of technology adoption on employment focusing on the general equilibrium effects in the labor market and labor polarization. Brynjolfsson and McAfee (2012) suggests that new digital technologies are having a structural impact on employment and are to blame for jobless growth. Frey and Osborne (2016) using data on occupations for the US labor market predict that about 47% of US jobs could be at risk due to computerization. Autor and Dorn (2013) suggest that the falling cost of automating routines and codifiable job tasks are one of the main determinants of the polarization of employment and wages in the US labor market. A common finding of this employment literature is the fact that new technologies, especially via automating routines, are having a strong impact on labor demand and the relative returns to different labor tasks and skill intensities.

In contrast, the empirical literature examining the direct impact of innovation on firm-level employment, while providing support to the existence of skill-biased technical change, has been more positive about the impact on labor (see Vivarelli (2012) and Calvino and Virgillito (2018)

for overviews and other articles in this volume).[4] Overall, this literature finds a positive direct link between product innovation and firm level employment. Harrison et al. (2014) find positive innovation to employment elasticities close to unity for France, Germany, Spain and the UK; while Crespi and Tacsir (2013) find also very similar results for Chile, Argentina, Costa Rica and Uruguay. A more puzzling result, however, in some of these studies is the fact that the effects of process innovation are ambiguous.

3 Methodology

3.1 Baseline specifications

To examine the impact of innovation on firm-level employment we adopt the empirical approach based on the model developed in Harrison, Jaumandreu, Mairesse, and Peters (2014). This model is well suited for cross-sectional data that contain information on a firm's current activities as well as its growth in sales and employment over a recent period. The crucial component of the model is the share of current sales due to newly introduced or improved products which is reported in our data, and that captures the extent of innovation. In this section, we briefly review key aspects of the methodology and our extensions to the model; we refer the reader to the Harrison et al. (2014) paper, and Crespi et al. (2018) in this volume for a detailed exposition.

Under this framework, the impact of innovation on employment depends on product innovation, process innovation and the relative efficiency parameters. The effect of product innovation on employment growth is represented by the difference in efficiency between the production processes of old and new products. When new products are produced more efficiently than old ones, output growth due to new products leads to smaller increases in employment compared to old products. The relative efficiency parameter could also be capturing to what extent product innovations are geared towards being more cost effective versus improving quality, which

[4] Some important studies include Harrison, Jaumandreu, Mairesse, and Peters (2014) for France, Germany, Spain, and the United Kingdom, Hall, Lotti, and Mairesse (2008) for Italy, Dachs and Peters (2014) for European countries and Benavente and Lauterbach (2008) for Chile. Also, see Castillo, Maffioli, Rojo, and Stucchi (2014), Brown, Earle, and Lup (2005) for effects of innovation policies on firm performance.

are arguably labor eroding and enhancing, respectively. While the relative efficiency parameter is firm specific, the country average could be correlated with the level of development of the country or the type of innovation introduced, given that firms in more advanced countries may be more capable to improve relative efficiency of new products.

While product innovations are not necessarily related to changes in the efficiency of production, process innovations are typically directed towards improvements in efficiency. A further complication arises when a firm introduces both product and process innovations simultaneously. A priori, it is difficult to predict the relative efficiency of *ex post* firm-level innovation and will depend on the combination of innovation outcomes and the degree of efficiency of the firm.

Following Harrison et al. (2014), our reduced form estimating equation is as follows:

$$l_i - g_{1i} = \alpha_0 + \alpha_1 d_i + \beta\, g_{2i} + u_i, \qquad (1)$$

where l stands for rate of employment growth over the period (i.e., between the year $t = 0$ and $t = 2$), g_1 and g_2 are corresponding rates of nominal sales growth for old and new products, and u is the unobserved random disturbance. The change in average efficiency in the production of old products is denoted as α_0 and d_i includes an additive component for process-only innovators; β measures the relative efficiency gains from the production of new products.

As discussed in Harrison et al. (2014), sales growth due to new products, g_2, can be correlated with the error term leading to a downward bias in β. This correlation can be driven by: 1) economic factors such as unobserved productivity shocks affecting both the decision to innovate and sales growth, and 2) measurement error when using nominal sales due to the lack of observed prices and appropriate price deflators. Instrumental variable (IV) methods are used to deal with the endogeneity stemming from the measurement issue while it is assumed that endogeneity effects from unobserved productivity shocks are more benign given timing differences. Variables that explain the success of the product innovation's sales but that are uncorrelated with its price differences relative to the old product should serve as good instruments. We use a series of indicator variables that measure whether the product innovation was geared towards extending the

market, whether the firm invests in R&D, and whether the innovation is completely new to the firm; these variables do not imply any necessary impact on prices. In all estimates, we evaluate and report the strength of our instruments as suggested in Stock and Yogo (2005) and conduct Sargan–Hansen over-identification tests.

3.2 Extensions to the baseline model

We consider some extensions to augment our baseline findings on the impact of different types of innovation on employment. First, firms can also introduce organizational innovations, which lead to change in the organizational structure of the firm such as departments or units within the firm, and that can also affect efficiency and the level of employment. Thus, we extend the original model in (1) to include the impact of organizational innovation on employment to also affect the relative or trend efficiency term.

$$l_i - y_{1i} = \alpha_0 + \alpha_1 d_i + \alpha_2 org_i + \beta y_{2i} + u_i, \qquad (2)$$

Second, the net impact of product and process innovation is not fully identified in (1), since the relative efficiency term in β reflects both types of innovation when implemented simultaneously. Section 3.3 reviews an approach to further disentangle the impacts of product and process innovation on employment.

A final extension of the model is related to the definition of process innovation. The way process innovation is defined in most innovation surveys using Oslo-manual guidelines includes any improvements on production or delivery methods, which can range substantially in their impact on efficiency and employment. Following the literature emphasizing the impact of automation on employment, we extend the model and decompose process innovation between innovations that imply some degree of automation in the production process and other types of process innovation.

3.3 Accounting for heterogeneity of impact

In our benchmark analysis, we pool the data and estimate equation (4) with controls for each country and sector. However, there are strong reasons to believe that unobserved heterogeneity can play a role at the country and sector level, given existing technology differences. In order to account for these differentiated effects, we extend the model to estimate random intercepts and coefficients for different clusters, defined by pairs of country and sectors (see Rabe-Hesketh and Skrondal, 2012, for more details on the random effects model).

Rewriting equation (1) as:

$$l_{ij} - g_{1ij} = \alpha_0 + \alpha_1 d_{ij} + \beta\, g_{2ij} + u_{ij}, \tag{3}$$

where i indexes over the firm and j represents a distinct industry within a particular country. For any two firms from the same country and sector, it might be unrealistic to assume that the residuals u_{ij} and $u_{i'j}$ are uncorrelated. We decompose the total residual or error, $u_{ij,}$ into a shared component between firms in the same country and sector group or cluster, ξ_{1j}, and a firm specific component, $\varepsilon_{ij,}$ that represents the within group residual – referred to as level-2 and level-1 residuals, respectively. Similarly, we can specify a country and sector specific random slope, ξ_{2j}, that affects g_2 in addition to the fixed component, β. As a result, we allow the intercept and slope to vary by each country's sector. The model becomes,

$$l_{ij} - g_{1ij} = (\alpha_0 + \xi_{1j}) + \alpha_1\, d_{ij} + (\beta + \xi_{2j})\, g_{2ij} + \varepsilon_{ij}, \tag{4}$$

where it is assumed that the random effects have zero means conditional on observables and the level-1 error term has zero mean given the covariates and random effects:

$$E(\varepsilon_{ij}|\mathbf{X_j},\xi_{\mathbf{1j}},\xi_{\mathbf{2j}}) = \mathbf{0} \tag{5}$$

Furthermore, given $\mathbf{X_j}$ the random intercept and random slope follow a bivariate distribution assumed to have zero mean and covariance matrix $\mathbf{\Psi}$.[5] The model is estimated via maximum likelihood with bootstrapped standard errors after obtaining predicted values of growth in new sales, g_2, in a first-stage equation as in the IV approach outlined in the previous section.[6]

The country and industry specific components, ξ_{ij}, represent the combined effects of omitted variables or unobserved heterogeneity at the country and industry level. Because of the shared components, the model accounts for within-country and sector dependence among the total residuals. The random intercepts and slopes can be interpreted as latent variables whose variance terms are estimated along with the other parameters and the variance term of the level 1 residual ε_{ij}. However, after estimating the model's parameters, including the random intercept and random slope variances as well as their covariance, $\widehat{\psi_{11}}, \widehat{\psi_{22}}, \widehat{\psi_{21}}$, we can obtain estimates of the random intercepts and slopes by an auxiliary regression that regresses the predicted total residuals on g_{2ij} and a constant term.[7]

The added advantage of exploiting within country and sector variation is that it allows us to try to disentangle the impact of different types of innovation measured by β in equation (4) discussed above. While the estimate of β is unbiased econometrically, to disentangle the true impact of product from process innovation on employment we would need to decompose this coefficient. We do so by exploiting the variance across the estimated country-sectors coefficients from the random-coefficient model, $\beta + \xi_{2j}$, which are regressed on the intensity of product, process and organizational innovation in each country-sector cluster. Thus, the following equation is estimated:

$$\beta_j = \alpha + \gamma \mathbf{w_j} + u_{1j}, \qquad (6)$$

where j denotes the country and sector, α is a constant term and $\mathbf{w_j}$ is a vector of innovation intensity measures representing the incidence of innovation outcomes in each country-sector group.

[5] The total residual is $\xi_{ij} \equiv \xi_{1j} + \xi_{2j}x_{ij} + \varepsilon_{ij}$ and its conditional variance is $Var(\xi_{ij}|\mathrm{x_{ij}}) = \psi_{11} + 2\psi_{21}x_{ij} + \psi_{22}\mathrm{x}_{ij2} + \theta$, where the level-1 residual is assumed homoskedastic and with conditional variance $Var(\varepsilon_{ij}|\mathrm{x_{ij}}, \psi_{ij}) = \theta$.

[6] The default case sets ψ_{21} and the corresponding correlation coefficient to zero but this assumption can be relaxed using the option covariance (unstructured).

[7] Alternatively, it can also be estimated using empirical Bayes as described in Rabe-Hesketh and Skrondal (2012).

Although imperfect, given the aggregation of the data to the country-sector level, this two-stage estimation allows us to explore the effect of combining different types of innovation with product innovation on the impact of sales from new products on employment.

4 The Data

In order to examine the innovative behavior of firms in developing countries, we use the World Bank 2013-2015 Enterprise Survey and its linked innovation modules. This is the most comprehensive set of cross-country surveys on innovation with the same sampling methodology and questionnaire carried out to date. The survey uses a stratified sampling strategy, where firms are stratified by industry, size, and location; and large and medium size firms tend to be over represented. An advantage of the survey is that it collects substantial balance sheet and other information regarding the investment climate, which enables the linkage of innovation efforts to performance and potential obstacles.[8] The mode of data collection is face-to-face interviews.

The innovation survey differentiates between product and process innovation, and two non-technological innovations, marketing and organization. However, there is significant confusion when identifying the different types of innovation outcomes by firms in the survey. For example, new marketing processes such as discounts, new packaging or new client segments are sometimes identified with process or product innovations. The fact that interviewees provide a recorded description of the product and process innovations allows us to verify the identified product and process innovations, and clean the wrongly attributed cases, or the cases that do not constitute an innovation at all (the detailed methodology to clean the data is described in the Cirera and Sabetti, 2016). Overall the cleaning exercise results in a significant decrease in both product and process innovation rates.[9]

[8] Sector breakdown is usually manufacturing, retail, and other services. For larger economies, specific manufacturing sub-sectors are selected as additional strata on the basis of employment, value-added, and total number of establishments' figures. Geographic regions within a country are selected based on which cities/regions collectively contain the majority of economic activity. Enterprise Surveys implemented in Eastern Europe and Central Asian countries are also known as Business Environment and Enterprise Performance Surveys (BEEPS) and are jointly conducted by the World Bank and the European Bank for Reconstruction and Development. For more details see Enterprise Surveys (http://www.enterprisesurveys.org).

[9] Overall both product and process innovation rates fall from 37% to 31%.

Our final data set consists of samples of firms from 53 countries in manufacturing and services across four major regions: Africa, Europe-Central Asia (ECA), Middle-East and North Africa (MENA), and South Asia. In total our estimation sample is based on pooled cross-sections totaling over 15,000 firms where sufficient information on innovation and employment is available.[10] Firms report both their sales and employment in the year the survey was conducted, which for most firms surveyed was in 2013 or 2014, as well as three years prior. In addition, firms report information about their innovation activities, such as product and process innovations but also including organization innovations, and the share of sales attributed to new innovations that allows us to decompose employment growth into its respective components driven by old products and new products growth.

Table 1 reports the number of firms in our sample across regions and by type of firm measured in terms of size, age and sector. Overall, we observe that the firms in our sample tend to be small, with fewer than 20 employees, and at the same time tend to be older, with the majority of firms operating for more than 10 years.

5 Results

5.1 Employment and sales growth among innovators and non-innovators

Table 2 compares firm employment and sales growth rates over the three-year period from the year in which the survey was completed, according to whether firms are non-innovators, process-innovators only, product innovators, and if so, whether they are also engaged in process innovation, by manufacturing and services sectors respectively. Innovation is larger in manufacturing than in services; and process and organizational innovation are the less prevalent types of innovation; although process innovation is more prevalent when combined with product innovation than organizational innovation. In manufacturing, roughly 46% of firms report undertaking some form

[10] DRC, Ghana, Kenya, Namibia, Nigeria, South Sudan, Sudan, Tanzania, Uganda, Zambia, Malawi, Bangladesh, India, Nepal, Pakistan, Albania, Armenia, Azerbaijan, Belarus, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Estonia, Georgia, Hungary, Kazakhstan, Kosovo, Kyrgyzstan, Latvia, Lithuania, Macedonia, Moldova, Mongolia, Montenegro, Poland, Romania, Russia, Serbia, Slovakia, Slovenia, Tajikistan, Turkey, Ukraine, Uzbekistan, Egypt, Israel, Jordan, Lebanon, Morocco, Tunisia, West Bank and Yemen.

of innovation, with roughly 27% of firms reporting product innovations. Slightly over half of firms with product innovations have also introduced process innovations.

The rate of technological innovation is strikingly large in Africa and South Asia. Rates for innovations defined as new to the national or international market are much lower, and the gap between new to the firm and new to the national market is significantly large for Africa and other low-income countries. Radical innovations - innovations new to the international market – and patenting are rare in these countries. Overall, this suggests large imitation rates in these regions, and the fact that many of the innovations that are new to the firm in most of the countries in our sample are likely to be marginally incremental; such as new product additions or small improvements.

On average, firms that have introduced product and/or process innovations tend to display higher growth rates both in employment and sales, although these differences are not statistically significant. Interestingly, when decomposing sales growth for product innovators into respective shares made up of old and new products, we observe that sales growth for old products tends to be negative. As a result, much of the overall sales growth for these firms tends to be driven by new products, which might be due to cannibalizing old products. We also find that the cumulative distribution functions of employment and sales growth for innovators, segmented by firm size, do not stochastically dominate those for non-innovators. However, we observe some evidence that innovating firms exhibit higher growth rates for parts of the distribution, particularly for larger firms.

Table 2 also highlights the challenge described in section 3 when trying to disentangle the employment effect of product and process innovation. Around half of product innovators in manufacturing and 30% in services also implement process innovation. In these cases, process innovations are also likely to affect the productivity terms of the new products introduced.

5.2 The impact of innovation on employment

5.2.1 The impact of innovation on employment growth

Table 3 shows the results of estimating equation (1) by OLS for the pooled sample and by different geographical and income regions, using process innovation only dummy, sector-country and size dummies.[i] Starting with the coefficient of process innovation only, this is negative and statistically significant, with the exception of ECA, MENA and high-income countries; suggesting that cases where firms only introduce new processes, increases in efficiency can result in a decrease in employment growth. Adding the process innovation only coefficient to the constant term allows retrieving the original intercept that reflects the trend productivity term of old products with negative sign.[11] The trend productivity parameter (-α) is negative for Africa, MENA and low-income countries sample, which suggests that labor productivity for all products in these regions have decreased. For the entire sample, the trend productivity parameter is positive for process innovators only, and also positive for all firms in South Asia, ECA and high-income countries.

The main coefficient of interest, the elasticity of sales attributed to product innovation on employment growth, is statistically significant and positive in all specifications, and 0.6 on average for the whole sample. Interestingly, the coefficient is below unity in all OLS specifications, which implies that new products are produced more efficiently than old products as in equation (2) and suggests a positive but less than proportional employment elasticity to innovative sales that results in some labor displacement. This elasticity is also larger in the MENA and South Asia region and middle-income countries in general; suggesting lower efficiency impacts arising from innovation.

However, as discussed in section 3, OLS estimates are likely to be biased due to the endogeneity of the sales growth of new products when the true real change in growth of sales is unobserved due to the lack of appropriate price deflators. To correct for this potential endogeneity, Table 4 shows the instrumental variables estimations. As instruments for sales growth from new products, we follow Harrison et al. (2014) and use variables that are correlated with the success of

[11] An alternative to splitting the sample by region and income groups would be to estimate the full sample and include interaction terms. However, the potential endogeneity of the innovation sales to employment changes described in previous sections makes it very difficult to find suitable instruments for the interactive terms.

the innovation but that do not necessarily imply any directional change in prices of new products relative to old ones; a dummy for whether the product is completely new to the firm and whether the firm invests in R&D. R&D investments and innovation decisions are made in advance, prior to any information on prices, and, therefore, unlikely to be correlated with the error term, while important to determine the success in sales terms of the innovation introduced. The Sargan-Hansen Over identification test (Stock and Yogo, 2005) p-value confirms the validity of the instruments used.

The new IV estimates suggest that the previous OLS estimates on the innovation elasticities were biased downwards. The new estimates are also statistically significant but much larger, especially in Africa and to a lesser extent in low and middle-income countries. This suggests that the employment growth associated to product innovations is larger in these lower income regions because of having lower efficiency gains from innovation in new and upgraded products. While these new or upgraded products are able to capture market share, the fact is that most innovation in these regions are marginal – new to the firm and simple upgrades – with little or no efficiency gains. The coefficient is significantly lower in the ECA region and high-income countries, suggesting that innovation in countries closer to the technological frontier increases employment less than proportionally driven by the larger impact of innovation on efficiency in this region.

While the size of the elasticities is conditional to the instruments used, the estimates for middle income-countries are similar to those in Crespi and Tacsir (2013) ranging from 1.16 for Argentina to 0.95 in Uruguay, and close to unitary elasticities. The estimates in Table 4 are, however, much smaller to our smaller sample of high-income countries than to Harrison et al. (2014) sample of EU countries that show unitary elasticities for manufacturing and in some countries for services as well.

In contrast, the impact of process innovation only in the instrumented specifications on employment growth is never statistically significant. In addition, regarding the trend productivity parameter of all products, in the instrumented specifications the parameters are not statistically significant, except for the ECA region where it is positive and marginally significantly negative for South Asia, MENA and low-income countries.

Given the fact that these results could be explained by the different sector composition in these regions, we examine how these elasticities vary across aggregate sectors, technological intensity levels and skill intensity. Table 5 examines the employment effects of product innovation according to whether the firm operates in the manufacturing or services sector. We find that β is on average higher in manufacturing than in services, but in both cases, we cannot reject that the elasticity is equal to one, suggesting that product innovations are not labor displacing in both sectors. When we disaggregate manufacturing into its degree of technological intensity, measured according to a sector's intensity of R&D expenditures, we find that the effect increases when comparing low-tech firms to medium-high tech firms, from 0.917 to 1.002; although again we cannot reject that the elasticities are equal to unity. Columns 5 and 6 of Table 5 display results from splitting the sample of manufacturing firms into those with high intensity of skilled employees and low intensity firms defined as having a share above or below the sector average, respectively. As expected, the coefficient is larger in low skill intensity firms than in high skill intensity firms, suggesting larger employment effects for firms with less skilled employees, again, explained but by lower efficiency gains from innovation in low skill sectors.

5.2.2 Employment effects of more novel types of innovations

Typically, product innovation range significantly from radical innovations that are new to national or international markets, to cases of marginal product upgrading or introductions of new product lines. It is possible, therefore, that the additional impact on employment from introducing product innovations largely depends on their degree of novelty, since the elasticity depends on relative efficiency; and the more different the new product introduced to existing products, the more likely relative efficiencies differ. On the one hand, more radical innovations may have a larger impact on relative efficiency and as a result on labor displacement. On the other hand, less novel innovations can have a marginal impact on quality attributes and product differentiation, and less so on relative efficiency.

Table 6 displays the results when considering innovation for differing degrees of innovation novelty. First, we consider separately when innovation is a product upgrade versus a completely new product. On the one hand, if the product upgrade is large, β in equation (1) - the ratio between

the efficiency parameter of the old product and the new product – is likely to depart from unity, but with unknown direction. On the other hand, small product upgrades are unlikely to change relative efficiency. In the case of new products, the size and direction of changes is difficult to predict. The estimates show a negative coefficient on product upgrade reducing the elasticity; however, both for manufacturing and services, the coefficient is not statistically significant. Therefore, we do not find different employment elasticities when comparing quality upgrade vs product diversification.

We extend the analysis to the degree of "novelty" on innovation and compare innovation defined as new to the firm vs innovations that are new to the national or international market in columns (3) and (4). Again, we should expect changes in relative efficiency correlated with the degree of novelty, and likely in some cases to greater relative efficiency, lowering β. The results suggest that in manufacturing there is no additional effect in terms of employment of more radical innovations as compared to new to the firm innovation - imitation. On the other hand, the results for services suggest a positive employment elasticity premium of more radical innovation. One potential explanation of this is the fact that knowledge transmission in service firms often relies on human and organizational factors more intensively than in manufacturing, where management plays a central coordinating role. This more intensive use of labor, especially skilled labor, can be exacerbated in more radical innovations and, therefore, more than offsetting any labor productivity changes (see Tether, 2003).

5.2.3 Process innovation and automation

Frey and Osborne (2016) discuss how automation and computerization of tasks may eventually render many labor tasks obsolete. One caveat of the previous estimates is the fact that the measure of process innovation mixes very diverse processes, ranging from the introduction of any new process or delivery methods with the introduction of an automated process (see Appendix). Thus, the process innovation dummy is likely to be capturing very different efficiency generating processes.

Table 7 reports estimates of the baseline instrumented model where we decompose process innovation into two components: innovations that involve automation of any process and non-

automation innovations.[12] The estimates suggest that the effects of automation only on employment growth in general are not statistically significant for most sector disaggregation. This suggest that while automation is likely to have significant effects on the skill and task composition of firms, at least in the short-run and when automation only is implemented as an innovation, it does not appear to have a direct impact on firm employment. An exception to this appears to be services where we find a marginal statistical negative effect.

5.2.4 Impact of organizational innovation

As suggested above, one important type of non-technological innovation that can have an impact on employment growth is organizational innovation. Changes in firms' departments and organizational structure, outsourcing of tasks or management structure changes (see Appendix for the definition) are likely to have an impact on employment growth. Table 8 shows the baseline estimates of including an organizational innovation only dummy. Organizational change was only reported in the survey for large and medium size firms, so there is a significant loss of observations when analyzing this type of innovation.

The estimates show that as it was the case for process innovation only, organizational innovation only does not seem to impact employment growth in the instrumented regression. Improvements in the organizational structure or outsourcing of tasks do not appear to affect the level of short-term employment of the firm when implemented in isolation from other types of innovation.

5.2.5 Disentangling the effects from different types of innovations

As discussed above, a large share of firms conducts both product and process innovation simultaneously. As a result, the elasticity on growth in sales due to new products is to some extent also capturing the impact of process innovation when implemented along with a product innovation. Given the nature of our data and the lack of within firm variation, it is difficult to isolate the true effects of each individual type of innovation. As suggested above and given the large heterogeneity of impact across sectors and countries, we estimate the model allowing the

[12] The size and country-sector dummies are restricted to add up to zero.

elasticities to vary by country and sector pairs j, and then in a second stage, estimate how these β_j are affected by the intensity of innovation in each country-sector; measured by the share of firms in the country-sector doing product and process innovation; organizational innovation and automation.

Table 9 shows the results of estimating the baseline model to obtain the random coefficients, including the estimated standard deviations of the random intercept and random slope and the residual standard deviation needed for estimating the these β_j. Column (1) reports the model with random intercept and slope, where we test whether the random slope is needed in addition to the random intercept. The likelihood ratio test where the null hypothesis $\xi_j = 0$ is strongly rejected in favor of the random intercept model. The average elasticity estimate of growth in sales of newly introduced or improved products is similar to the results of the fixed effects model in Table 4, where we obtained an estimate of 0.938. Based on the estimated variance, roughly 95 % of random slopes fall in the interval ($0.962 \pm 1.96 \times \sqrt{0.243}$). Column (2) implements the same model but includes the effect of organizational innovation only. Column (3) decomposes process innovation into automation and process innovation excluding automation.

Once these β_j for all country sectors are estimated, Table 10 shows the estimates of the second stage, where we use the variation of these β_j across country-sectors and analyze their correlation with the share of product innovators in the country-sector, the share of product and process innovators, product and organizational innovators, and product and automation innovators. In each model, we include a specification that controls for sector specific effects that may explain differences in the elasticity of employment growth with respect to the growth in new products, for example the degree of sector capital and labor intensity. The estimates for the share of product innovation across specifications is not statistically significant, suggesting that having more innovators in the country-sector does not affect the size elasticity. The coefficient on product and process innovation is as expected negative in column (1), suggesting that country-sectors that more intensively combine product and process innovation have lower elasticities and potential efficiency gains from process innovation that reduces the additionality on employment creation. However, when we control for sector effects in column (2) the coefficient becomes statistically not significant. Overall, our estimates find no impact of the intensity of process innovators on the

elasticity of sales of new products on employment, not even when combined with product innovation.

A more interesting picture emerges when using the intensity of product innovation jointly with automation. In this case, we find that the employment elasticity of new sales due to innovation is lower in sectors where product innovations are more likely to be accompanied by automation of production processes. Therefore, while automation alone may not affect short-term employment, when accompanied with product innovation is likely to increase efficiency and reduce the ability of generating additional employment.

6 Conclusion

This paper has analyzed empirically the impact of firm-level innovation on employment. The main result of this paper is that product innovation, when successful in bringing additional sales to the firm, has a positive direct impact on employment in the short-run. The extent to which sales cause additional employment, however, is directly related to the impact on efficiency resulting from the innovation process, which seems to vary with the income level of the country. The empirical estimates suggest that in lower and middle-income countries, and especially in Africa, where innovations are more incremental and there may be less efficiency gains due to the innovation, the employment growth associated with a dollar increase in sales from innovative products is larger than in middle and high-income countries. In fact, the model estimates predict that for most countries if all products could be replaced by new or upgraded products, the overall level of employment of the firm would be at least as high as the previous level. This is an important finding given some pessimistic views that innovation efforts are often entirely labor saving for the firm. On the other hand, for high income countries, new sales attributable to innovation generate new employment but at lower rates since the new or upgraded products are more efficient in the use of labor than old products, and, therefore, generating some labor displacement.

The findings point towards product innovation as the main channel of employment creation. Organizational innovation does not appear to have any impacts on employment changes, when considered alone or when implemented with product innovation. The same occurs with

process innovation, which does not seem to impact employment, even when considering the introduction of automation processes. It is likely that the main effects of these types of innovation are on the type of labor - skill biased technical change - rather than the quantity of labor. However, we find some support to the idea that automation may actually displace labor by reducing the employment elasticity of product innovation when these are introduced jointly.

The implication of these results for policy are important. Innovation policy, when effective in generating additionality on successful innovation activities, even via imitation, can also be an important policy to increase employment in the short-run. This is especially the case for those countries farthest away from the technological frontier, where the effect on employment from generating new sales due to innovation is largest. On the other hand, for higher income countries, the additionality impact on employment is lower given their greater ability to generate productivity gains in new products.

More work is, however, needed to better understand the short-run impact of innovation on the skill composition of the firm and the effect of the different types of innovation combined. While innovation is likely to be skill biased, we know very little about the extent of potential skill displacement within the firm, on the job learning and more generally, what the impact on unskilled workers inside innovative firms is.

Table 1: Sample description

		All	Africa	South Asia	ECA	MENA
Total		15033	1652	3322	6658	3401
Size	Small (< 20)	0.54	0.73	0.36	0.61	0.49
	Medium (20-99)	0.36	0.23	0.47	0.33	0.38
	Large (100+)	0.10	0.04	0.17	0.06	0.12
Age	< 5	0.05	0.13	0.04	0.05	0.01
	5 to 9	0.22	0.30	0.19	0.24	0.16
	10 to 14	0.21	0.21	0.19	0.24	0.18
	15 to 19	0.19	0.12	0.17	0.23	0.18
	20 +	0.32	0.22	0.41	0.23	0.46
Industry	Food	0.08	0.09	0.09	0.06	0.12
	Textiles	0.04	0.02	0.07	0.02	0.04
	Garments	0.06	0.04	0.07	0.05	0.10
	Wood, Paper	0.03	0.03	0.04	0.03	0.02
	Publishing, Printing	0.03	0.03	0.02	0.03	0.03
	Chemicals	0.03	0.02	0.06	0.02	0.03
	Plastics	0.03	0.01	0.07	0.03	0.02
	Non metallic mineral products	0.03	0.02	0.04	0.04	0.01
	Basic metals, products	0.07	0.07	0.10	0.05	0.07
	Machinery	0.03	0.01	0.07	0.03	0.00
	Electronics	0.04	0.01	0.10	0.03	0.01
	Furniture	0.03	0.07	0.03	0.03	0.03
	Construction	0.05	0.04	0.02	0.08	0.03
	Motor vehicle services	0.03	0.05	0.02	0.03	0.02
	Transportation	0.05	0.03	0.03	0.05	0.06
	Wholesale	0.09	0.09	0.02	0.12	0.08
	Retail	0.18	0.24	0.07	0.24	0.12
	Hotels & Restaurants	0.06	0.12	0.05	0.04	0.06
	IT, Professional services	0.02	0.01	0.02	0.02	0.01

Note: Based on estimation sample covering 53 countries. *Source:* Enterprise Surveys 2013-2015.

Table 2: Growth of employment and sales, innovators and non-innovators

	Manufacturing					Services				
	All	Africa	South Asia	ECA	MENA	All	Africa	South Asia	ECA	MENA
No of Firms	7846	691	2541	2665	1949	7187	961	781	3993	1452
Non-innovators (%)	0.54	0.49	0.26	0.68	0.74	0.74	0.59	0.45	0.79	0.84
Process-innovators only (%)	0.18	0.23	0.31	0.12	0.09	0.10	0.22	0.23	0.07	0.07
Product-innovators	0.27	0.28	0.43	0.20	0.16	0.16	0.20	0.31	0.15	0.10
of which product & process innovators	0.51	0.52	0.60	0.40	0.37	0.29	0.40	0.26	0.26	0.29
Employment growth										
All firms	0.05	0.08	0.07	0.05	-0.01	0.06	0.10	0.08	0.06	0.04
Non-innovators	0.03	0.07	0.05	0.05	-0.02	0.05	0.08	0.09	0.05	0.04
Process innovators only	0.07	0.06	0.08	0.08	0.00	0.08	0.12	0.07	0.09	0.02
Product innovators	0.06	0.11	0.08	0.04	0.01	0.08	0.12	0.08	0.07	0.07
Sales growth (Nominal)										
All firms	0.09	-0.04	0.16	0.16	-0.05	0.07	-0.18	0.16	0.16	-0.07
Non-innovators	0.06	-0.06	0.14	0.15	-0.06	0.07	-0.13	0.17	0.16	-0.07
Process-innovators only	0.11	-0.05	0.15	0.18	-0.06	0.05	-0.12	0.15	0.17	-0.10
Product innovators	0.13	-0.02	0.17	0.16	-0.01	0.05	-0.37	0.16	0.15	-0.01
of which: Old products	-0.15	-0.33	-0.12	-0.05	-0.34	-0.20	-0.70	-0.13	-0.05	-0.28
New product	0.30	0.28	0.34	0.25	0.28	0.26	0.23	0.36	0.24	0.25

Note: Employment and sales growth are measured over three-year period from time survey was completed. Employment growth is measured as change in full-time employees. Sales growth is measured as change in local nominal currency. For product innovators, we measure overall sales growth, sales growth due to old products, and the share of current period sales attributed to New products. *Source:* Enterprise Surveys 2013-2015.

Table 3: Effects of innovation on employment (OLS), by Region and Income Categories

	All Countries	by Region				by Income		
		Africa	South Asia	ECA	MENA	Low	Middle	High
Process innovation only	-0.037***	-0.104**	-0.051***	-0.003	0.026	-0.067**	-0.027**	0.005
	(0.01)	(0.04)	(0.02)	(0.02)	(0.03)	(0.03)	(0.01)	(0.02)
Sales growth in new products	0.640***	0.373***	0.686***	0.597***	0.817***	0.583***	0.756***	0.479***
	(0.04)	(0.13)	(0.06)	(0.06)	(0.09)	(0.10)	(0.04)	(0.08)
Size 20 to 99	0.008	-0.026	0.010	0.022*	-0.011	-0.019	0.014	0.016
	(0.01)	(0.04)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Size 100 and over	0.010	-0.036	0.031*	-0.005	-0.002	-0.012	0.015	0.017
	(0.01)	(0.09)	(0.02)	(0.02)	(0.02)	(0.04)	(0.01)	(0.02)
Constant	0.016**	0.282***	-0.038**	-0.061***	0.076***	0.113***	0.004	-0.036***
	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
N	14688	1756	3322	6657	2953	3468	6024	5196

Note: Note: Coefficients and standard errors robust to heteroskedasticity and 1, 5, and 10 percent levels of significance are denoted by ***, ** and *, respectively. All regressions include country-industry and size dummies, constrained to sum to zero. Dependent variable is net labor growth (minus growth in sales of old product). *Source:* Enterprise Surveys 2013-2015.

Table 4: Effects of innovation on employment (IV), by Region and Income Categories

	All Countries	by Region				by Income		
		Africa	South Asia	ECA	MENA	Low	Middle	High
Process innovation only	-0.001	-0.000	0.004	0.007	0.034	-0.024	0.022	0.022
	(0.01)	(0.06)	(0.04)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)
Sales growth in new products	0.938***	1.726***	0.945***	0.803***	1.030***	1.013***	1.056***	0.797***
	(0.07)	(0.52)	(0.20)	(0.08)	(0.12)	(0.25)	(0.10)	(0.10)
Size 20 to 99	0.006	-0.043	0.007	0.021*	-0.013	-0.022	0.012	0.013
	(0.01)	(0.04)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Size 100 and over	0.005	-0.077	0.029	-0.008	-0.008	-0.019	0.012	0.011
	(0.01)	(0.09)	(0.02)	(0.02)	(0.03)	(0.04)	(0.02)	(0.03)
Constant	-0.004	0.243*	0.208*	-0.133**	0.229*	0.172**	-0.039	-0.086
	(0.01)	(0.15)	(0.12)	(0.06)	(0.13)	(0.09)	(0.08)	(0.08)
N	14688	1756	3322	6657	2953	3468	6024	5196
R-Squared	0.29	0.32	0.19	0.19	0.14	0.39	0.18	0.18
First stage F-statistic	404.59	29.35	25.57	268.94	126.34	56.89	136.92	218.77
Sargan–Hansen test p-value	0.87	0.93	0.44	0.58	0.84	0.64	0.86	0.49

Note: Coefficients and standard errors robust to heteroskedasticity and 1, 5, and 10 percent levels of significance are denoted by ***, ** and *, respectively. All regressions include country-industry and size dummies, constrained to sum to zero. Dependent variable is net labor growth (minus growth in sales of old product). Instrumental variables are indicator whether new product is *completely new to the firm* and whether firms *invest in R&D*. *Source:* Enterprise Surveys 2013-2015.

Table 5: Effects of innovation on employment (IV), by Sector and Technological Intensity

	Manufacturing	Low Tech	Medium Tech	Medium High Tech	High Share of Skilled Workers	Low Share of Skilled Workers	Services
Process innovation only	0.008	0.022	-0.019	-0.007	-0.007	0.042	-0.007
	(0.02)	(0.02)	(0.05)	(0.06)	(0.02)	(0.04)	(0.02)
Sales growth in new products	0.966***	0.917***	0.894***	1.002***	0.838***	1.189***	0.914***
	(0.11)	(0.13)	(0.24)	(0.28)	(0.14)	(0.18)	(0.09)
Size 20 to 99	0.003	-0.004	0.009	-0.004	0.017	-0.012	0.012
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Size 100 and over	0.012	0.022	-0.020	0.008	0.010	0.015	-0.013
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
Constant	0.033	0.190**	-0.073	0.011	0.112*	-0.180*	-0.006
	(0.05)	(0.09)	(0.09)	(0.08)	(0.06)	(0.10)	(0.09)
N	7424	3875	1995	1554	4784	2640	7118
R-Squared	0.23	0.25	0.27	0.23	0.28	0.27	0.34
First stage F-statistic	144.34	104.88	25.20	18.26	80.13	48.14	286.28
Sargan–Hansen test p-value	0.90	0.40	0.66	0.14	0.74	0.31	0.60

Note: Coefficients and standard errors robust to heteroskedasticity and 1, 5, and 10 percent levels of significance are denoted by ***, ** and *, respectively. All regressions include country-industry and size dummies, constrained to sum to zero. Dependent variable is net labor growth (minus growth in sales of old product). Instrumental variables are indicator whether new product is *completely new to the firm* and whether firms *invests in R&D*. High and Low Share of skilled workers represent firms with share of skilled workers above or below the population mean in the most recent fiscal year, respectively. *Source:* Enterprise Surveys 2013-2015.

Table 6: Effects of innovation on employment (IV), by Product Innovation Novelty

	Product Upgrade vs. Completely New		New to Firm vs New to National or International market	
	Manufacturing	Services	Manufacturing	Services
Process Innovation only	0.004	-0.007	0.013	-0.015
	(0.01)	(0.02)	(0.02)	(0.02)
Sales growth in new products	0.963***	0.918***	1.031***	0.742***
	(0.06)	(0.08)	(0.11)	(0.11)
Product Upgrade × Sales growth in new products	-0.043	0.100		
	(0.07)	(0.10)		
New to National or Intl × Sales growth in new products			-0.135	0.331***
			(0.10)	(0.12)
Constant	0.039	0.059	0.036	0.063
	(0.05)	(0.09)	(0.05)	(0.09)
N	7424	7264	7424	7264
R-Squared	0.23	0.33	0.23	0.35

Note: Coefficients and standard errors robust to heteroskedasticity and 1, 5, and 10 percent levels of significance are denoted by ***, ** and *, respectively. All regressions include country-industry and size dummies, constrained to sum to zero. Dependent variable is net labor growth (minus growth in sales of old product). Instrumental variables are indicator whether new product is *completely new to the firm* and whether firms *invests in R&D*. *Source:* Enterprise Surveys 2013-2015.

Table 7: Process Automation

	All Firms	Manufacturing	Low Tech	Medium Tech	High Tech	Services
Automation only	0.001	0.023	0.02	0.009	0.029	-0.106*
	(0.02)	(0.03)	(0.07)	(0.06)	(0.04)	(0.05)
Process only (excl. automation)	-0.001	0.003	-0.02	-0.032	0.02	-0.002
	(0.01)	(0.02)	(0.07)	(0.05)	(0.03)	(0.02)
Sales growth in new products	0.941***	0.974***	1.002***	0.898***	0.922***	0.899***
	(0.07)	(0.11)	(0.27)	(0.24)	(0.13)	(0.09)
Constant	-0.002	0.037	0.012	-0.075	0.194**	0.062
	(0.01)	(0.05)	(0.08)	(0.09)	(0.09)	(0.09)
N	14688	7424	1554	1995	3875	7264

Note: Coefficients and standard errors robust to heteroskedasticity and 1, 5, and 10 percent levels of significance are denoted by ***, ** and *, respectively. All regressions include country-industry and size dummies, constrained to sum to zero. Dependent variable is net labor growth (minus growth in sales of old product). Instrumental variables are indicator whether new product is *completely new to the firm* and whether firms *invests in R&D*. High and Low Share of skilled workers represent firms with share of skilled workers above or below the population mean in the most recent fiscal year, respectively. *Source:* Enterprise Surveys 2013-2015.

Table 8: Effects of innovation on employment, accounting for Organizational Innovation

	OLS	IV
Process innovation only	-0.062***	0.021
	(0.02)	(0.03)
Organizational only	-0.068***	0.013
	(0.02)	(0.03)
Sales growth in new products	0.621***	1.118***
	(0.05)	(0.12)
Size 20 to 99	-0.003	-0.006
	(0.01)	(0.02)
Size over 100	0.000	-0.004
	(0.000)	(0.02)
Constant	0.074***	0.070
	(0.01)	(0.08)
N	4404	4404
R-Squared	0.35	0.31
First stage F-statistic		109.43
Sargan–Hansen test p-value		0.34

Note: Coefficients and standard errors robust to heteroskedasticity and 1, 5, and 10 percent levels of significance are denoted by ***, ** and *, respectively. All regressions include country-industry and size dummies, constrained to sum to zero. Dependent variable is net labor growth (minus growth in sales of old product). Instrumental variables are indicator whether new product is *completely new to the firm* and whether firms *invests in R&D*. *Source:* Enterprise Surveys 2013-2015.

Table 9: Effects of innovation on employment, baseline model to compute random coefficients

	Product, Process	Product, Process, Organizational	Product, Process, Automation
Fixed part			
Process Innovation Only	-0.007	-0.001	0.000
	(0.01)	(0.03)	(0.02)
Organizational only		-0.004	
		(0.03)	
Automation only			-0.033*
			(0.02)
Growth new sales	0.962***	1.023***	0.961***
	(0.08)	(0.12)	(0.09)
Size 20 to 99	-0.038***	-0.080***	-0.038***
	(0.01)	(0.02)	(0.01)
Size over 100	-0.063***	-0.109***	-0.063***
	(0.01)	(0.02)	(0.01)
Constant	0.032***	0.058**	0.032***
	(0.01)	(0.03)	(0.01)
Random Part			
ψ_{11}	0.038***	0.090***	0.048***
	(0.01)	(0.02)	(0.01)
ψ_{22}	0.243***	0.198***	0.242***
	(0.01)	(0.02)	(0.01)
θ	0.432***	0.413***	0.432***
	(0.00)	(0.01)	(0.00)
N	11762	3696	11762
Pseudo log-likelihood	-7154.07	-2114.88	-7153.14
LR test	0.00	0.00	0.00

Note: Coefficients and standard errors robust to heteroskedasticity and 1, 5, and 10 percent levels of significance are denoted by ***, ** and *, respectively. Standard errors computed using 5,000 bootstrapped replications. Dependent variable is net labor growth (minus growth in sales of old product). Instrumental variables are indicator whether new product is *completely new to the firm* and whether firms *invests in R&D*. Based on multilevel mixed-effects linear regression with random intercept and slope for g_2 within country and sector. Sample restricted to country sector clusters with at least 10 observations. *Source:* Enterprise Surveys 2013-2015.

Table 10: Regressions of Random Coefficients on Innovation intensities

	Product & Process		Automation		Organizational	
	(1)	(2)	(3)	(4)	(5)	(6)
Share product innovation	0.045**	0.041*	0.043	0.036	0.018	0.023
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)
Share product and process innovation	-0.025	-0.007	0.078	0.101	-0.026	-0.032
	(0.03)	(0.04)	(0.07)	(0.07)	(0.02)	(0.03)
Share product and automation			-0.244**	-0.244**		
			(0.11)	(0.11)		
Share product and organizational					-0.028	-0.027
					(0.02)	(0.02)
Constant	0.955***	0.951***	0.952***	0.948***	1.026***	1.019***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
N	306	306	306	306	284	284
Sector dummy	No	Yes	No	Yes	No	Yes
R-Squared	0.02	0.07	0.05	0.10	0.01	0.05
Pseudo log-likelihood	588.81	596.65	521.88	529.38	369.92	376.08

Note: Dependent variable in all regressions is random slope parameters for country/sector clusters derived from (4). Independent variables are respective innovation rates – share of innovators in each country/sector cluster. Coefficients and standard errors robust to heteroskedasticity and 1, 5, and 10 percent levels of significance are denoted by ***, ** and *, respectively. Standard errors computed using 5,000 bootstrapped replications. Based on 306 country sectors with at least 10 observations. *Source:* Enterprise Surveys 2013-2015.

Effects of Innovation and Financing on Startup Firm Survival and Growth *

Alexandru Minea[†] Leonard Sabetti[‡]

Marcel Voia[§]

Draft

Abstract

The launch of new business ventures is an important source of dynamism for both advanced and transitioning economies. Startups that survive initial years (including a small subset that exhibit high-growth) contribute in aggregate to economic employment and productivity gains. Using a detailed survey dataset with information on firm strategy, financing, innovation activities and founder characteristics, we study the performance in terms of survival and growth of a representative cohort of American startup firms launched in 2004 over an eight-year period until 2011; overlapping with the recession of 2008-2009. Our results highlight the differential effects of initial financing on survival depending on the business cycle and the role of the firm's technological intensity in lowering hazard rates. Rich measures on firm innovation activities related to product innovation are key determinants of firm growth. We also investigate the role of innovation in securing external financing as a potential mechanism for early stage firm survival and growth.

Key Words: new firm entry, duration models, entrepreneurial finance, innovation, entrepreneurship.

JEL Classification: C41, C14, D21, L60.

*Acknowledgements: We are grateful for helpful comments from seminar participants at CEA, CERDI and Stata Users workshop; especially Patrick Plane and Nestor Sawadogo. Disclaimer: Certain data included herein are derived from the Kauffman Firm Survey. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Ewing Marion Kauffman Foundation.

†CERDI, Universite´ Clermont Auvergne (France) Email: alexandru.minea@uca.fr

‡CERDI, Universite´ Clermont Auvergne (France) Email: leonard.sabetti@etu.udamail.fr

§Laboratoire d'Économie d'Orléans (LEO) Faculté de droit, d'économie et de gestion Université d'Orléans (France) Email: marcel.voia@univ-orleans.fr

1 Introduction

Increasing availability of longitudinal datasets have uncovered large and persistent heterogeneity in the performance of firms despite the apparent stability in aggregate macro measures; see Dosi, Lechevalier, and Secchi (2010) for an overview. Firm dynamics and entrepreneurship literature have documented how mature, large firms tend to stagnate whereas new, young firms tend to grow very quickly conditional on survival, resulting in an important source of reallocation and aggregate productivity growth for both advanced and transitioning economies; Haltiwanger, Jarmin, and Miranda (2010), Foster, Haltiwanger, and Krizan (2002), Brown and Earle (2010). However, Decker, Haltiwanger, Jarmin, and Miranda (2016) have recently documented a decline in business dynamics, or firm turnover, driven by declining business start-up rates and a decreasing role of dynamic young firms in the economy since the onset of the Great Recession. Because of their impact on impact broader aggregate variables, the performance of new or early stage firms remains an important public policy issue.

In this paper, we use a novel dataset that tracks a representative sample of firms from a cohort of US-based startups launched in 2004. A major strand of the literature on early stage firms focuses on the role of finance, such as credit constraints and sources of funding via debt or equity. Another strand highlights the role of the founder, human capital, firm strategy and innovation. In this paper, we incorporate both sets of factors with an emphasis on the role of financial variables and the degree of innovation of the firm in studying startup firm performance. Whereas traditional measures of innovation have focused on R&D expenditures and patents, a third source of indicators increasingly made available in surveys emphasizes firm innovation activities such as the successful introduction of different types of innovation to the market. Further, traditional measures for innovation based on R&D expenditures and patents may not be reflective of innovation activities, particularly for newly established firms or firms far from the technological frontier. As a result, we are able to investigate to what extent these richer measures that capture firm innovation activities relate to and affect startup firm performance. Our contribution to the literature stems from deriving new empirical evidence on the importance of innovation activities for a sample of startups observed over their early life-cycle in the United States. Previous studies using the same dataset did not include information on innovation activities which were made available in the survey only

in subsequent years. Studies on the role of innovation activities have typically focused on mature established firms where the data are often cross-sectional in nature given the challenge of repeat sampling.[1] Further, these data sometimes lack a complete picture of the firm with less detailed measures of financing and characteristics of the founder.

In our first set of results, we study the factors that drive firm survival, with a particular focus on initial financing as well as the degree of firm innovativeness. While financing affects firm survival ex post, lenders also incorporate their estimate of firm survival – a measure of firm quality – into their lending decisions. In our sample of firms, only a subset launch with formal external debt financing as many firms are initially financed with owner equity or debt. Given the timeframe under study, we are also able to shed light on the effect of the Great Recession in heightening firm exits or shutdowns. We find that a firm's initial financial position characterized by a greater reliance on formal debt financing positively affected survival in normal times, but the effect reverses during the crisis. To study the effect of firm innovativeness, we use measures of the firm's industrial sector's degree of technological intensity as well as whether it performs R&D; which also correlate with proximity to the technological frontier. We find strong statistical differences in the hazard rates for firms in the high-tech sector which peak mid-way during the observed life-cycle.

Second, we carry out a set of regressions that explore the link between firm level innovation and growth, conditional on our sample of surviving firms; growth measures proxy for firm performance. Our measures of innovation are expanded to include whether the firm brings product innovations to market (outputs). Measures on innovation outputs become available in the survey beginning in 2009 – the firm's sixth year of operation. Applying a quasi difference-in-difference framework, we assess the impact of innovation activity occurring through the firm's observed lifespan on performance from a before-after perspective exploiting the longitudinal nature of our data. For robustness, we test for selection on initial conditions and examine effects on growth in the pre-treatment period; given a lack of suitable IVs and an experimental setting.

Finally, a potential mechanism explaining the relationship between firm innovation and growth lies in the role between financing and innovation strategy. A firm's attempt to innovate and build competitive advantage potentially influences its ability to apply for and secure additional

[1]Examples include Community Innovation Surveys (CIS) for Europe and Business Research and Development and Innovation Survey (BRDIS) or Annual Survey of Entrepreneurs (ASE) for the US.

rounds of external financing over time; simultaneously affecting and driving growth. Alternatively, firm innovation activities reveal underlying characteristics of the entrepreneur in terms of his or her motivation or degree of confidence which could also influence financing decisions. In a series of regressions, we evaluate to what extent firm innovation measures explain the firm's ability to raise additional external debt financing over its lifecycle.

The rest of the paper continues as follows. Section II provides an overview of the literature on firm survival. Section III describes the data and Section IV reviews our empirical methodology. Section V discusses our results and Section VI concludes.

2 Connection to the literature

2.1 The role of finance

A distinguishing feature of young, startup firms is the opaqueness and information asymmetry surrounding their future business plans and survival prospects, Robb and Robinson (2014). The extent to which young, startup firms are unable to access sufficient financial capital and the resulting effects on performance and survival remain open questions. The effect of financing constraints can operate through various channels. In seminal papers, Stiglitz and Weiss (1981) showed that banks ration credit due to information asymmetries and Berger and Udell (1992) documented evidence of credit constraints using data from survey of loan officers. Evans and Jovanovic (1989) test whether entrepreneurs are more likely to be wealthy and in their theoretical model some projects become unprofitable for a financially constrained entrepreneur. Cabral and Mata (2003) show how much of the distribution of firm size evolution can be accounted for by initial startup financial constraints.

The literature has also emphasized the importance of firm startup financial capital and leverage. Does capital structure drive firm performance or does firm performance drive capital structure? Zingales (1998) studied this issue in a period of deregulation in the trucking industry and showed how some inefficient firms survive due to deep pockets. A central prediction from the finance literature posits that entrepreneurs will rely on a greater share of debt financing to the extent that their business venture entails a greater degree of certainty. Myers and Majluf (1984) predicted a pecking order theory where firms prefer to finance investment beginning with internal funds, then debt and then equity so as to minimize adverse selection. Zetlin-Jones and Shourideh (2012) find

that most investment financed by privately held firms is financed through borrowing. However, greater leverage may adversely affect a firm's chances of survival, as discussed in Zingales (1998). Firms may eventually suffer from debt overhang, unable to finance new projects, or, they may fall victim to predation by deep-pocketed competitors. In contrast, high leverage could work in favor of firm's survival by forcing early restructuring. Furthermore, high leverage may indicate a firm's aggressive expansion strategy. Finally, firms that launch with ample initial financing may also be more likely to misallocate that capital compared to leaner startups that expand in a more cautious, incremental manner.

Using data on corporate firms, Spaliara and Tsoukas (2013) find that survival is more sensitive to financial indicators during a period of crisis as compared to more tranquil times. To the extent that the recession was triggered by a financial crisis as opposed to traditional demand or supply side factors, the direction of causality should flow from the firm's financial position to its performance and survival. Young firms with a greater reliance on formal outside debt or short-term liquidity provisions would be more adversely affected in the event that their banking relationship deteriorated or future access to funding was reduced. Furthermore, to the extent that the policy response was aimed towards larger financial institutions, which do not typically specialize in small business lending, suggest that the financial crisis' adverse effect on young, startup firms may have been amplified. Using the same data as in this paper, Berger, Cerqueiro, and Penas (2014) find that a greater small bank geographical presence improves the survival prospects of young firms in normal times, but that this effect disappeared during the financial crisis.

Finally, a historical feature of the US banking sector is its presence of a large number of small, local financial institutions which have been found to specialize in forming lending relationships with small and young businesses. For example, Petersen and Rajan (2002) find that distance matters for provision of funds to small firms (though declining with improvements in information technology). A priori, the depth of the banking sector in the US might suggest that credit markets have largely filled the niche of providing finance to startups, however the evidence is mixed, and many argue that a market failure (still) exists. Brown and Earle (2013) have found positive effects on job creation from SBA loan programs.

2.2 The role of innovation

Without innovation, models of firm dynamics typically predict that firms stagnate after reaching a certain size. A large empirical literature mainly using data for OECD countries documents a robust positive relationship between firm-level innovation and productivity (Hall, 2011); innovation and employment (Harrison et al., 2008), as well as some innovation inputs, such as R&D, and productivity (Hall et al., 2011). Arrighetti and Vivarelli (1999) and Vivarelli and Audretsch (1998) find that more innovative new entrant firms tend to enjoy superior post-entry performance.

To measure innovation, one can focus on both measuring inputs and innovation activities, and/or measuring innovation outcomes. The early innovation measurement literature focused on a specific set of innovation inputs that were easier to quantify, for instance R&D, or the intensity of the technology used. These early efforts were followed by the implementation of the Oslo Manual type of surveys, which mainly focus on measuring innovation outcomes such as product/process improvements or patents at the firm level. Under the Oslo manual, innovation is principally defined as whether the firm introduces a new product to market, a measure which is self-reported. In contrast, firms in the high-tech sector are defined based on whether their ISIC industry has a larger than average share of R&D, or higher shares of employment in STEM fields. See Cirera and Maloney (2017) for a more detailed discussion. Mairesse and Mohnen (2010) discuss how qualitative questions in surveys can provide information on firm innovation activities useful for bench-marking and contributing to evidence-based policy; further, they review econometric methods to deal with the qualitative nature of responses.

3 The Kauffman Firm Survey

Almost a half a million startup firms are launched annually in the United States. Yet, little comprehensive data exists that allow researchers to study firms at their most interesting stage: initial years from launch. Existing databases, such as Compustat and Amadeus, track mature, established firms and generally do not contain information beyond that obtained from publicly available financial statements. Some past surveys of startup firms in the US are Survey of Small Business Finances and Panel Study of Entrepreneurial Dynamics I and II.

To bridge the gap, the Ewing Marion Kaufman Foundation commissioned a longitudinal study of new businesses in the US, known as the Kauffman Firm Survey (KFS). The survey follows 4,928 firms that launched in 2004 annually until 2011. The survey questionnaire contains detailed information on the firm, including industry, physical location, employment, profits, intellectual property, business strategy, and financial capital, as well as information on business owners, including age, gender, race, ethnicity, education, previous industry experience, and previous start-up experience. The initial survey design called for 5,000 interviews, with a target of 3,000 interviews for high-technology businesses given particular interest in these firms among researchers. While data is collected on firm location, the small samples at the geographic level, such as by county, limit the ability to relate variation in local economic conditions to firm level outcomes.

The sampling frame for the panel of business startups was created using Dun & Bradstreet's (D&B) database of business establishments that started in 2004 in the United States, which totaled roughly two hundred and fifty-thousand firms. D&B maintains a large commercial database of businesses compiled through various public and industry sources. Year of launch is defined as the first year the business began operations and took steps to incorporate itself; as a result, a firm's first year may not always entail sales or hiring employees posing a challenge for accurately measuring firm age.

In order to obtain a larger sample of startups in high-technology fields, the data was partitioned into strata according to industrial technology categories, based on a classification scheme developed by Hadlock, Heckler, and Gannon (1991). Table A.1 provides the final classifications of high and medium technology businesses, determined according to each industry's respective share of employment in research and development (R&D) using data from the BLS Occupational Employment Statistics program and based on three-digit level Standard Industry Classification (SIC) code. DesRoches, Barton, Ballou, Potter, Zhao, Santos, and Sebastian (2007) provide additional details.

Figure 1 displays the number of new establishments born as well as jobs created by establishments less than 1 year old annually since 1994; data is from the Business Employment Dynamics (BED) compiled from records from a federal-state cooperative program known as Quarterly Census of Employment and Wages (QCEW). While firm-level data can be compiled by aggregating establishments under common ownership using employer tax identification numbers, almost all

new establishments with fewer than 500 employees represent new firms. Figure 2 compares survival rates between establishments in the BED with firms in the KFS. Survival rates for cohorts of new business establishments are displayed by year of birth. While some of these new establishments may not coincide with new firms, the survival rates track fairly closely the survival rates for the firms launched in 2004 in the KFS. To improve the comparison, the survival rates for KFS firms are conditioned on firms with positive employment in the startup year. Unlike the KFS, there are no sampling error issues with the BED data. Whereas shutdown for BED is defined as establishments that revert to no employment for four consecutive quarters, firms in the KFS are deemed to shutdown based on direct reporting to the annual follow up survey.

Figure 3 plots the number of employees across the entire cohort of startup firms in our sample over time separately for all firms and surviving firms only. The number of total employees climbs rapidly in 2005 but begins to fall in 2006 as firms exit. In contrast, the total number of employees across surviving firms exhibits a more steady, gradual increase over time.

3.1 Overview of Firms in KFS

A challenge of empirical studies of entrepreneurship is distinguishing between types of entrepreneurs. Using data from the Statistics of the U.S. Business (SUSB) compiled by the U.S. Census Bureau, Hurst and Pugsley (2011) finds that most small businesses have little intention to grow or innovate. More recently, Brown, Earle, Kim, and Lee (2019) find higher levels of innovation in foreign-born compared to US-born entrepreneurs in high-tech sectors using data from the US Annual Survey of Entrepreneurs. An advantage of the KFS is its oversampling of high-technology firms, which also are likely to be high-growth innovative firms, in contrast to typical startups that eventually fall in the small business category. According to Table A.2, entrepreneurs in technology industries are more likely to have higher levels of human capital, to invest in R&D and to introduce a new product to market. Interestingly, however, on average they tend to start up at a smaller size, as measured by total assets.

However, to situate our study relative to popular interest in startups, we emphasize our focus is on the median, or typical, new firm compared to that of the high growth startup that is in the right tail of the distribution and which is sometimes borne out of incubators or VC financed. Out of our sample of 4000 firms, fewer than 1 percent obtain any form of outside equity over the observed

eight-year period. Almost all launch from the founder's home, are financed principally by its founder(s) and less than half launch with outside debt (formal bank issued debt).

Table A.3 provides a count of firm exit by year throughout the time period of study. By 2011 over half of the firms in the initial baseline year report shutting down their business, which is consistent with empirical estimates on new firm survival. Table A.4 breaks down the number of firms by two-digit industry according to the NAICS classification. Professional services, manufacturing and the retail sector make up the three largest industries.

3.2 Financing Decisions

The standard life-cycle theory of small firm finance assumes that young startups initially rely on informal channels of credit followed by increasingly formal sources as the firm establishes its business history. Robb and Robinson (2014) finds that the sample of startup firms in the KFS display a heavy reliance on formal debt financing, in line with the notion that startups also seek financing where capital is more readily available, Cosh, Cumming, and Hughes (2009). Similarly, we investigate firms' financial decisions based on both the type (debt vs. equity) and source (informal vs. formal). Capital can be provided by the business owner, an insider (family, friend), or an outside (formal lending institution or venture capital, angel financing).

While we focus on the source of finance, an important distinction raised by Robb and Robinson (2014) is the role of risk-bearing vs. liquidity. For instance, a business loan might be obtained via the founder pledging personal assets in the form of collateral, and as a result, the loan serves as an equity-like instrument for the founder. Personal bank loans obtained by the owner are classified as outside debt. Table 1 provides an overview of the mean levels of initial financing by type. We observe that the two main sources of financing are owner equity and outside debt. While roughly 40% of firms report some level of outside debt, the average amount is considerably larger than that reported by owner equity. In other words, outside debt is important for firms which obtain it. Owner debt plays a role for some firms, but the average reported amount is small. A handful of firms do report receiving large sums of financing via outside equity. Overall, we observe that outside financing is more reliant on debt, while inside financing is composed mostly of equity.

As a potential mechanism to understand firm exit, Table A.5 provides the evolution of average firm capital injections over time for the three most popular sources of finance. Initially in

2005 and 2006, the average amount of new capital obtained is similar between firms that survive and those that eventually fail. However, beginning in 2007, the average capital injections for exiting firms begins to drastically decrease, suggesting a precursor to firm exit is the inability to raise funds.

3.3 Firm-level innovation activities

In addition to sector level differences between firms in terms of intensity of technological inputs, the survey introduced questions beginning in calendar year 2009, or in the sixth wave of data collection, that asked firms about their innovation activities; namely, whether the firm introduced any new or significantly improved product or service and whether the innovation was new to the firm or new to the market (regional/national). From the later, we obtain a measure of novelty that distinguishes from imitation. Conditional on surviving firms, Table 2 provides a breakdown of reported firm-level innovation in the year it occurs by sector technological intensity. Further, it reports shares of firms that engage in R&D.

We observe overall large rates in innovation, patenting and R&D for firms in the medium and high-tech sectors. Most innovation activity occurred in 2009, potentially due to censoring as this was also the first year the question was introduced in the survey. Our initial prior is that firms launching product innovations sooner are more likely to display positive effects on growth measures. Further, not all firms introduce innovations, even in the medium or high-tech sector, and not all product innovations are potentially driven by R&D. Similarly, some firms invest in R&D but have yet to introduce any product innovation. In summary, Table 2 provides summary statistics of our set of innovation measures that form part of our regression approach in the next section.

4 Empirical Methodology & Hypotheses

As described in the previous section, the Kauffman Firm survey tracks a cohort of startup firms that launch in 2004 until 2011. We first study firm-level survival prospects over the period using a workhorse duration model. Second, we estimate a series of growth regressions that condition on the sample of surviving firms up to 2011; including the predicted survival rates as an explanatory variable. Our growth measures capture performance of the firm over an eight-year time span which

includes the Great Recession of 2008-2009. While we study separately survival from growth, they are inextricably linked as pointed out by Phillips and Kirchhoff (1989). In their study using longitudinal microdata on new establishments from the time period 1976-1986, the authors point out how firms that exhibit positive growth, particularly early on, have higher survival prospects; however, survival alone also tends to lead to (or increase chances of) future growth as most new establishments grow little in the first four years. As a result, our approach can be interpreted as studying firm survival as a measure of firm quality while firm growth as a measure of performance (or intensive margin).

4.1 Survival

Empirical studies on firm survival typically find negative duration dependence, the longer a firm operates the more likely it is to survive. This is consistent with a model where firms learn about their competitiveness as they age. Mature, older firms have very low failure rates. However, survival dynamics for in the early years of a firm's startup is more complex. Initially, a firm may undergo a "honeymoon" period for several years where business failure is infrequent. Eventually, firms uncover adverse market conditions, or face unsustainable financing, and firm exit peaks. Thereafter, firm exit rates decrease steadily with age; see Parker (2009), Huynh, Petrunia, and Voia (2012a), Huynh, Petrunia, and Voia (2012b), and Huynh and Petrunia (2010). Under this hypothesis, startup firm hazard rates should follow an inverse U-shaped pattern. Duration models are used extensively in the literature to estimate the hazard rate, or instantaneous probability of exit, for firms. Unlike logit models, they can account for right-censoring which can be important when observing firms for shorter time periods. See Cameron and Trivedi (2005) for a more complete discussion. The workhorse duration model is the semi-parametric (Cox) proportional hazards model, defined as:

$$h(t) = h_0(t)\varphi(X, \beta) \tag{1}$$

where $h_0(t)$ is called the baseline hazard and is a function of t only with all covariates set to zero. $h(t)$ represents the rate of failure at time t given that a firm survived in $t - 1$. $\varphi(X, \beta)$ can be interpreted as a scaling factor, and does not depend on t, typically specified in exponential form $exp()$ (Alternatively, accelerated failure time models allow the covariates to affect the hazard multiplicatively). The advantage of this approach is that the parameters β can be estimated,

or identified, without explicitly modeling the functional form of the baseline hazard (via partial likelihood). An important assumption in Cox models is the proportionality assumption, that is the effect from a covariate on the baseline hazard must be proportional over time, or in other words invariant. When the PH test fails, the suggested approach is to include the variable that failed the PH test as well as an interaction term with a time variable in a new regression.

We do not include time-varying firm specific covariates to avoid any feedback or endogeneity issues with respect to survival and as a result, the firm-level covariates are based on a firm's initial conditions. However, we include a time-period dummy variable to capture the role of the Great Recession on firms' hazard rates. The crisis dummy is constructed to take the value of 1 in years 2009-2009 and the value 0 otherwise. We interact this variable with the firm initial financial conditions and our set of innovation measures.

Our empirical approach is not immune from potential omitted variable bias, or unobserved heterogeneity; see for instance, Huynh and Voia (2017). The role of the entrepreneur in firm performance is typically void in economic stud- ies given the difficulty in observing measures of ability and talent. As a result, even a model with a comprehensive set of observables may fail to fully explain the variability in the data. One way to address the issue of unobserved firm-specific attributes is to employ (shared) frailty models, which allow for the presence of a latent multiplicative effect on the hazard function. While firm specific factors are important in duration models, unobserved heterogeneity can lead to misrepresentation of the overall firm hazard rate. When thinking about the role of the entrepreneur, observing information that captures ability and talent is typically sparse. Hurst and Pugsley (2011) find that many small business owners, including a share of young startups, do not have ambitions of growing their business or innovating.

One aspect of the firm that is important for understanding outcomes and that typically is unobserved is its degree of innovativeness. We do not include our measures on innovation activities (whether the firm introduces a new product) in our benchmark survival regression as these survey questions are only introduced in 2009. However, we use whether the firm is in the high-tech sector and whether it performs R&D as proxies for being close to the technological frontier (R&D intensive). Subsequently, we relate the probability a firm exits in 2010 and 2011 to firm innovation activities conditional on our sample in 2009.

Additional explanatory variables included in our benchmark survival model specification

control for firm-specific characteristics as well as local factors, according to the available data and informed by theory. Business characteristics (industry dummies, business location, legal status), founder's demographics (education, gender) and initial financing conditions, such as startup capital (total debt + equity), and leverage (measured as the ratio of outside debt to total debt and equity); see Huynh, Petrunia and Voia (2010). We use number of employees as well as startup capital to proxy for initial firm size. While the cohort of firms in our study all begin with the same, we include the founder's number of years of work experience. Work experience has been found in the literature to play an important role in explaining unobserved factors in firm performance. An entrepreneur with more work experience might possess greater knowledge or be better able to evaluate a potential business opportunity. Furthermore, work experience might also correlate with the founder's net worth, and ability to raise external finance.

4.2 Growth

In the previous section we study survival prospects of our cohort of startup firms over the 2004-2012 period conditional on initial conditions and degree of firm level innovation measured by sector R&D intensity. In this section, we condition our sample to firms that survive until 2011 and evaluate their performance from year of birth using a series of reduced form OLS or binary out- come regressions. In these set of estimates, we are able to exploit additional survey questions that were added to the survey beginning in 2009 related to firm innovation activities for each calendar year; outlined in Table 2. Specifically, these questions ask the firm whether they introduced any new and improved product or service and its degree of novelty, such as whether it is new to the firm or a regional market. As a result, we investigate the extent to which innovation outputs positively affects the growth of the firm beyond measures such as R&D and the firm's technology sector. Be- cause the firm reports introducing product innovations beginning in 2009, roughly midway through its observed lifespan, our model setup could be viewed as applying a quasi diff-in-diff framework where the dependent variable is measured in terms of a growth rate pre and post intervention; see Ci, Galdo, Voia, and Worswick (2015). Our measures of firm growth are in terms of end of period observations relative to startup year for 1) revenues, 2) number of full-time employees as well as 3) whether the firm experiences positive employment growth (binary). We estimate models separately using growth rates derived from the pre-intervention period (2004-2008), post-intervention period

(2008 to 2011) and over the entire observed lifecycle (2004 to 2011). In this way, growth regressions from the pre-intervention period allow us to examine parallel trends prior to the introduction of firm innovation reported in the survey as of 2009.

In these growth regressions, our set of predictors are time-invariant features of the firm and characteristics of the founder as well as initial financing and size conditions. Our treatment variables are the year in which the firm reports introducing a product innovation as of 2009. The sample of firms that report innovations may launch with more resources or unobserved factors that drive a firm to innovate would also affect performance simultaneously. While panel data models (or a quasi-experimental setting) could alleviate some of the endogeneity by differencing out omitted factors, our limited time frame and because the impact from time of innovation could span multiple years limit their appeal. Further, the available set of time varying predictors are largely limited to the evolution of firm financing which are also endogenous. To provide some indication of whether selection may be occurring, we estimate a set of regressions that evaluate whether innovative firms are larger or obtain greater levels of initial financing at year of birth. Our estimates can be interpreted in a reduced form way that measure the average impact after accounting for other available factors and conditional on survival. The contribution from our findings stems from providing evidence on startup firm performance using information on firm-related innovation activities that are largely unavailable in other data sources.

4.3 Interaction effects between financing and innovation

As discussed previously, finance plays a key role in affecting firm outcomes. In our survival and growth analysis, we evaluated the role of initial financial conditions as well as the role of firm innovation measures. However, the ability of a startup to raise additional financing over its life-cycle is a key enabler for its survival and growth prospects. One potential mechanism whereby firm innovation activities affect firm performance is through their indirect impact on the extent to which new, young firms are successful in raising external financing. To test this hypothesis, we undertake a series of regressions to explain whether the firm applied for new business loans, the outcome of such applications (approved or denied), and the level of additional external debt raised (intensive margin).

5 Results

5.1 Survival

Column 1 of Table 4 presents the results from our benchmark model. A coefficient of δ can be interpreted as raising or lowering the hazard by a factor of $1 + \delta$. Figure 4 evaluates the fit of our benchmark model relative to the empirical hazard. Overall, the estimated hazard tracks the empirical hazard closely, and we find that the semi-parametric Cox does a better job than a less flexible parametric model. However, we are less able to fit the data for 2011 due to censoring. The baseline hazard provides the results of the model assuming all coefficients are set to zero, which can be interpreted as accounting for the overall risk of firm exit not due to any firm-specific factor.

Overall, the explanatory variables have expected signs. Work experience, credit riskiness of the firm, R&D are all significant at the one percent level. Initial firm size measured by number of employees also displays highly statistical effects in lowering the hazard rate; confirming the interplay between survival and growth or size. A higher initial outside debt ratio lowers the firm hazard rate, but the initial overall level of financial capital has no effect. However, the effects of the financial variables are exacerbated when interacted with the economic crisis and play a lesser role before and after (or in normal times). A higher level of the outside debt ratio lowers the firm's hazard rate during normal times, but negatively affects the hazard rate during the economic crisis. The estimated coefficient switches from 0.716 to 1.423 in column 2. For initial startup capital, the effects are intensified when interacted with the economic crisis displaying coefficients above 1, but otherwise display muted effects not significantly different than zero in normal times. For example, the effect on firms with startup capital in excess of $100,000 switches from 0.892 to 1.4 in column 2. The results are consistent in further set of models with expanded set of predictors.

The finding that firms with greater levels of initial startup capital were more adversely affected during the Great Recession might be surprising or counter-intuitive at first glance but is in line with findings from the literature. In a study of banking crises and industry financial dependence, Kroszner, Laeven, and Klingebiel (2007) find that financially dependent firms grow faster in normal times but are hit harder in crisis times. The Great Recession proved to be an especially turbulent and uncertain time, and business plans that otherwise seemed a priori reasonable may have turned out to be unrealistic considering the shock to the economy and financial system. In

contrast, firms that raised smaller levels of initial financing were more likely to grow in stages and raise additional capital as their economic prospects improved, without overextending their business at the same time. Similarly, in a study using an unbalanced panel of 36,500 French startups from the period 1994-2000, Bonnet and Cressy (2016) find that banks incorporate an estimate of firm survival in lending decisions; while credit rationing raises failure rates, incorrectly over-extending credit could have similar effects.

With respect to our measures of firm innovation, Figure 5 displays firm hazard rates separately for firms in the high-tech sector compared to all others based on a split sample estimation. We observe an inverted U-shaped hazard rate for high tech startups where the hazard rate peaks in 2008; but an increasing hazard rate up until 2011 for all other firms. Given the empirical hazard rate for high-tech firms reaches its peak sooner suggests that these firms more quickly learn about their demand and survival prospects, as predicted by the theoretical literature; alternatively, firms in the high-tech sector were more able to weather the recession. Based on the survival models in Table 4, we find that R&D has a strong statistical effect on lowering the hazard; in contrast, there is no effect from patents. For robustness, we also interact these variables with the crisis variable in column 3 but do not find any effects.

Finally, we attempt to exploit the measures of firm innovation activities that are introduced into the survey in 2009. Given censoring, we were unable to introduce it in the survival model directly as firms exit prior to when the question was introduced. However, we find that firms that report introducing new products or service to market as of 2009 display a ten-percentage point decrease in their probability of exit in subsequent years (2010 and 2011); see Table 5 for marginal effects from a probit model. As a robustness check, we included estimated propensity scores from a Probit model on product innovation reported in Table 3 as additional predictors in the survival model (results available) but such an approach suffers from generated regressor bias. Overall, our results provide evidence on the role of innovation measures as important sources of unobserved heterogeneity when unaccounted for in firm survival models.

5.2 Growth

In this section, we present results from a set of regressions that evaluate the link between firm innovation and the resulting effects on firm performance over the time period 2004-2011. Performance

is measured in terms of growth in log revenues, change in the number of employees and whether the firm displays positive employment growth. Because many firms launch with no revenue in the first year, we use the first year where such measures become positive and adjust the growth rate based on the number of years until 2011. Our estimation sample comprises roughly 2,000 firms that survive until 2011. Our main set of innovation related treatment variables that we are interested in are measures of product innovation reported in calendar years 2009 to 2011. Further, we classify product innovation according to three types: overall product innovation (new to the firm or regional market), product innovation that is new to the market, and product innovation that coincides with R&D. In this way, we attempt to isolate differential effects according to the degree of novelty. Other set of innovation measures that we include are R&D, patent activities, and whether the firm is in the high-tech sector. We use a similar set of controls based on initial conditions of the firm, characteristics of its principal founder, along with state and industry effects. Regressions are implemented using cross-sectional survey weights.

Prior to evaluating growth regressions, Table 6 investigates firm initial conditions and whether firms that eventually engage in product innovation start off larger or obtaining greater levels of initial financing. We find only weak evidence that these firms have more employees at startup (coefficient magnitude of 0.393) and no evidence for initial revenues. We find a statistical effect at the five percent level, however, for initial levels of financing (log 0.3) which points to a potential channel whereby innovation interacts with financing; innovative firms are more likely to obtain outside financing which feeds into firm performance measured by size (labor and sales). In contrast, we find strong statistical negative effects for firms in the medium or high-tech sector on initial levels of financing.

Table 7 presents our set of growth regressions for sales and employment (where we model employment in terms of change in levels and whether growth is positive) according to three periods: full period (2004-2011), pre-treatment (2004-2008), and post-treatment (2008-2011). Growth results for the entire observed lifecycle show stronger statistical effects from product innovation on revenues than employment and the largest impacts occur when product innovation occurs simultaneously with R&D. In the pre-treatment period, we do not find any effects from product innovation on revenues and only some modest positive effect on employment growth; suggesting that the effect from product innovation is more causal in nature than selection. Growth results

for the post-intervention period are similar to the full-time period results but lower in magnitude; which could be driven by the Great Recession and the subsequent recovery.

Interpreting our results suffer from two empirical challenges. First, given that these set of innovation survey questions were introduced in 2009, it is possible that reported innovations may have in fact occurred sooner which proves to be a limitation in our analysis. Second, a time lag to impact from introducing a new product to market on firm outcomes likely holds given the nature of innovation. In addition, fewer firms report introducing innovations later in their observed lifecycle in 2010 and 2011 creating a censoring problem as we only observe growth up until 2011.

Table 8 displays the effects from select other predictors we include in our growth models based on the full period. We observe a low R^2 of 0.065 for the revenue growth regression highlighting the limited explanatory power of our observed variables. We do not find any effect from R&D or firms in the high to medium tech sector for any measures of growth while patents display a negative effect for employment growth. Both previous startup experience and years of work experience have no effect or are possibly negative. Years of experience correlates with founder age and the literature finds that successful startups are sometimes launched by younger entrepreneurs as well as older ones with greater experience depending on the industry and context; pointing to the overall difficulty of any predictor in having direct effects that hold on average. Among the financial variables, firms in the highest category of initial startup capital display statistically positive effects for all our growth measures. Initial levels of employment tend to be positively correlated with employment growth suggesting the role of persistence with initial startup size; however, we do not find an effect on revenue growth. Overall, in terms of economic magnitude of our coefficient estimates, we observe that the effects from our product innovation measures are roughly similar or larger in magnitude than effects from measures such as initial financing.

5.3 Interaction effects between innovation and financing

To investigate further the potential mechanism from the role of finance, Table 9 relates the log annual outside debt financing the firm raises over its lifespan post entry to its initial conditions and whether the firm innovates in subsequent years. We estimate regressions for the entire sample of firms with controls for firm exit by year as well as for the sample of surviving firms separately. Here we find statistical effects at the one percent level for firms that report innovation in 2009 (log

0.632 to 0.785) of roughly 15% increase relative to the mean; however, the magnitude of the effects is lowered when conditioning on the novelty of innovation possibly pointing to the riskiness embedded in innovation. With respect to other effects, there is some persistence as firms that obtain higher initial levels of financing also receive greater injections of outside debt in subsequent years on average. We observe that patents exhibit a marginally negative effect possibly in line with the literature that finds patent valuations to be highly skewed.

In Tables 10 and 11 we estimate a series of probability models for whether a firm reports applying for a business loan post entry and whether any business loan applications were denied (employing a Heckman selection model). Similarly, regressions are estimated for the entire sample with year of firm exit and separately for the sample of surviving firms. Firm size and initial financing positively affect the probability of loan applications. Perhaps due to endogeneity, outside debt ratio lowers the probability of loan denial (and the effect is stronger conditional on the sample of surviving firms). In contrast, credit risk increases the probability of loan denial but has no effect for the sample of surviving firms. Interestingly, when we examine the effects from our measures of innovation activities, we observe a positive effect on loan applications but little to no effect on loan denial.

6 Conclusion

Conventional wisdom holds that young, startup firms face poor survival prospects; let alone the prospect for growth and contributing to broader economic gains. However, firm turnover - and the role of successful startups - play an important role in driving productivity and employment growth. Small businesses - startups that survive but remain small - in aggregate are an important source of local employment and economic output; even if they are less efficient than their larger peers. In this paper, we use a longitudinal representative panel of US based start-up firms for the period 2004-2011 to investigate the drivers of firm performance in terms of survival and growth.

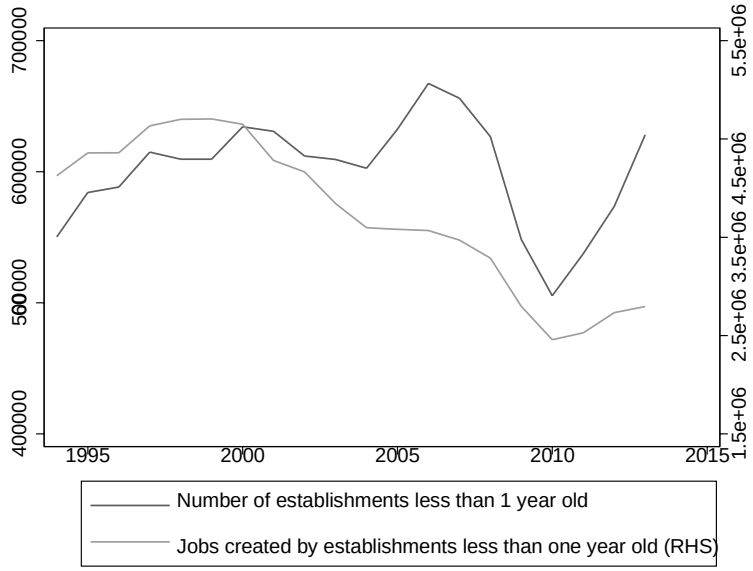
Our results focus on the role of initial financing conditions and measures that capture the degree of firm innovation. For firm survival, we find mixed results from the role of initial financing. Raising outside debt signals a high-quality firm; but too much debt can drive exit particularly during a recession. Conditional on survival, initial financing levels affect firm growth positively.

Further, many startups rely on personal/internal funds or personal debt.

While financing can influence firm outcomes, the nature of the firm in terms of its competitive strategy and degree of human capital also matter. Our results highlight the founder's work experience as a positive contributor to survival. Further, we measure the degree of firm innovation according to traditional measures derived from R&D intensity or whether the firm is in the high-tech sector. We find that high-tech firms display an inverted u-shaped hazard rate; compared to non-tech firms which display an increasing hazard rate during the time period under study. Further, we also find positive effects on survival using direct measures of reported firm innovation activities, such as whether a firm introduces a new product or service. We also find differential effects on firm growth based on these innovation measures; highlighting their importance in uncovering heterogeneity. Finally, we investigate a potential causal mechanism whereby more innovative firms have a higher probability of raising additional finance in their initial years; a crucial time for firm survival and growth.

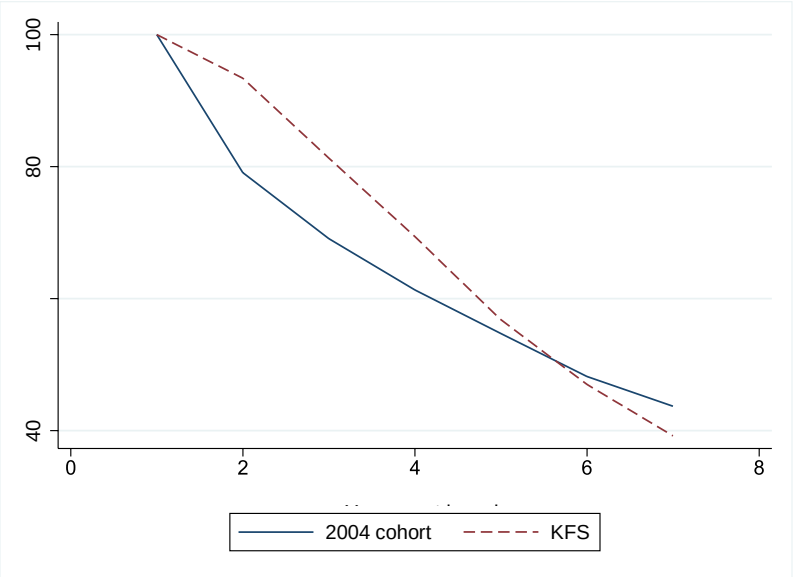
Our findings contribute to the literature on startups and entrepreneurship using a novel dataset overlapping with the business cycle around the Great Recession of 2008-2009. Our results are informative for policymakers in terms of highlighting not only the importance of financing but also the role of firm quality proxied by innovation measures. Promoting human capital and knowledge diffusion and better understanding their underlying determinants could help drive entrepreneurial businesses that both support local economies and compete on the global stage.

Figure 1: Business Employment Dynamics



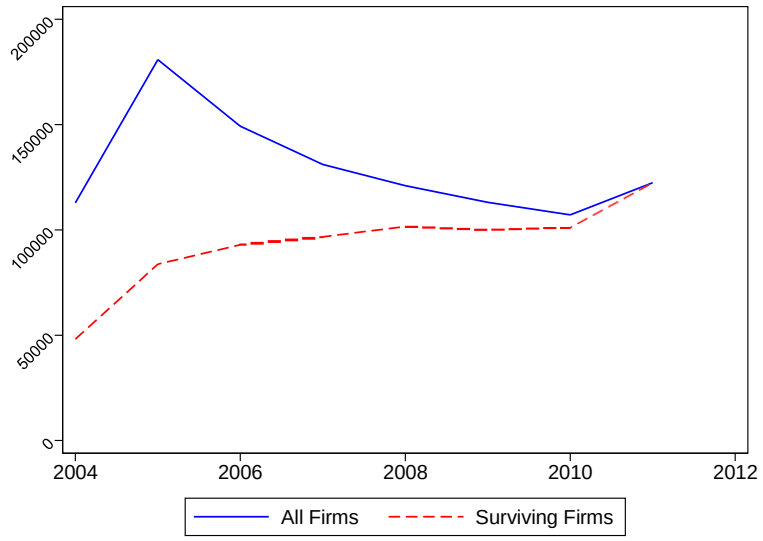
Source: Bureau of Labor Statistics Business Employment Dynamics (BED)

Figure 2: Survival Rates of Establishments and Firms in KFS



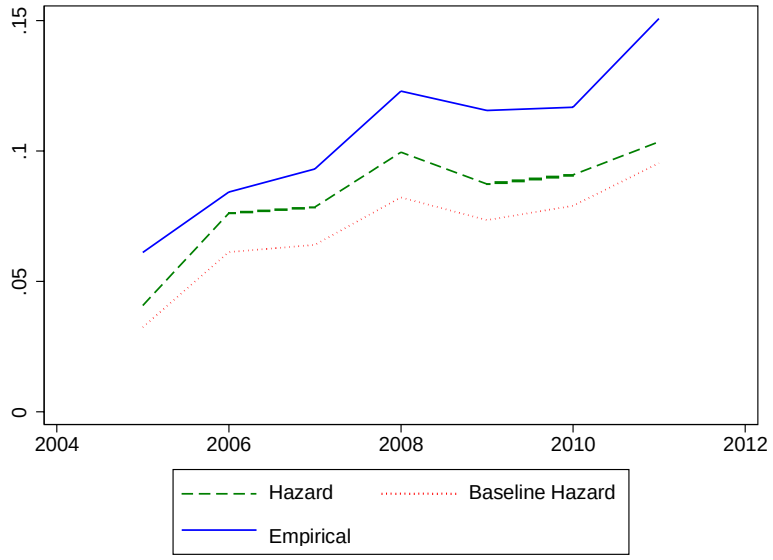
Note: Year of birth for new establishments in BED defined by positive employment for the first time in the database, while failure is defined by no employment in four consecutive quarters. Employment changes are measured from the third month of the previous quarter to the third month of the current quarter.

Figure 3: Total Stock of Employees from Entering 2004 Cohort of Startup Firms



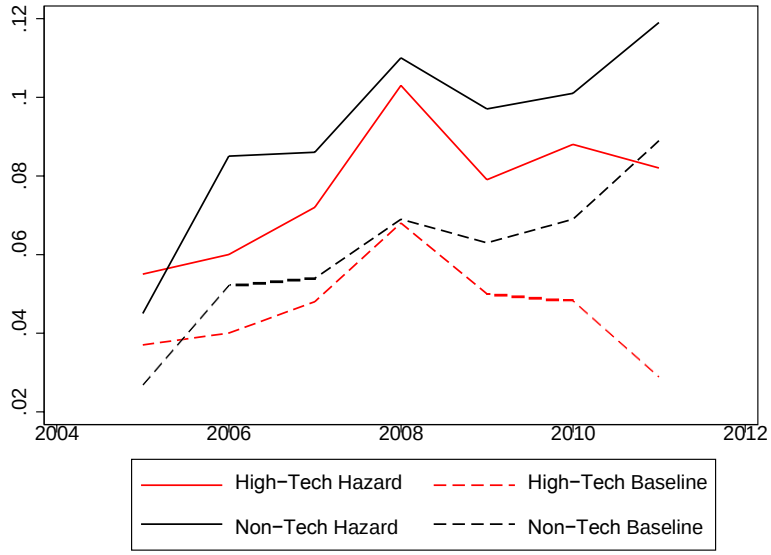
Note: Based on weighted count of 62,452 startup firms in Kauffman Firm Survey.

Figure 4: Hazards



Note: Baseline and hazard rates based on benchmark model. Empirical hazard based on raw data count.

Figure 5: Hazards: High-Tech vs. Non-Tech



Note: Baseline and hazard rates based on benchmark model. Empirical hazard based on raw data count.

Table 1: Sources of Financing for 2004 startups

	All Firms	Mean if > 0	Count	Survive	Fail
Owner Equity	29,188	38,951	3789	25,898	26,819
Owner Debt	2,592	10,441	1473	2,400	2,690
Inside Equity	579	43,865	215	264	698
Inside Debt	2,512	43,179	545	2,452	1,993
Outside Equity	3,655	550,496	267	2,725	3,576
Outside Debt	33,416	120,406	1814	39,299	28,352
Total Capital	91,646	105,887	4340	91,226	87,617

Note: Based on sample of 4,216 firms in Kauffman Firm Survey. Numbers displayed are in average dollar terms.

Table 2: Innovation rates

	High Tech	All Other
Product innovation - New to the firm		
2009	22.7	17.2
2010	9.9	8.8
2011	7.8	6.4
Cumulative	40.4	32.4
Product innovation - New to the market		
2009	12.7	11.0
2010	5.4	3.6
2011	3.5	3.0
Cumulative	21.6	17.6
Product innovation - backed by R&D		
2009	11.5	7.3
2010	3.6	1.8
2011	2.6	1.6
Cumulative	17.7	10.7
Patents in First year	5.90	2.00
Performs R&D	25.7	17.3

Note: Based on conditional sample of 2,007 surviving firms in Kauffman Firm Survey. All numbers are sample weighted and in percentage terms.

Table 3: Probability of product innovation, 2009

	New to the firm	New to the market
Bachelor degree	-0.006	0.044**
	0.02	0.02
Graduate degree	0.019	0.063***
	0.02	0.02
Number of past businesses	0.001	-0.002
	0	0
Years work experience	0.007	-0.003
	0.01	0.01
0 < Employees < 5	0.029	0.034*
	0.02	0.02
5 < Employees < 10	0.053	0.066**
	0.03	0.03
Employees > 10	0.081**	0.071**
	0.04	0.03
Outside debt ratio	-0.005	-0.025
	0.03	0.03
5,000 < Startup capital < 10,000	0.044	0.009
	0.03	0.03
10,000 < Startup capital < 25,000	0.075***	0.050**
	0.03	0.02
25,000 < Startup capital < 100,000	0.067**	0.031
	0.03	0.02
Startup capital > 100,000	0.088***	0.057**
	0.03	0.03
High credit risk	0.018	0.001
	0.04	0.04
High medium tech	0.037	0.021
	0.02	0.02
Patents	0.002	0.047
	0.05	0.04
R & D	0.310***	0.197***
	0.02	0.02
N	2337	2337
aic	2683	2048
ll	-1255.3	-967.5

Note: Probit marginal effects reported. The dependent variable is a dummy equal to one if the firm reports introducing product innovation in 2009, and zero otherwise. *, **, *** denotes statistical significance at the 10, 5 and 1 % levels respectively. Other founder demographics, business characteristics, state and sector controls not shown.

Table 4: Hazard Model of Startup Firm Survival: 2004 to 2011

		Model 1	**Model** 2	**Model** 3
Firm owner characteristics	Male	0.897**	0.896"*	0.896**
		0.05	0.05	0.05
	Bachelor degree	0.963	0.963	0.963
		0.06	0.06	0.06
	Graduate degree	0.918	0.918	0.918
		0.06	0.06	0.06
	Number of past businesses	1.005	1.005	1.005
		0.01	0.01	0.01
	Years work experience	0.905***	0.905***	0.905***
		0.02	0.02	0.02
Firm size and financing	0 < Employees < 5	0.827***	0.827***	0.827***
		0.05	0.05	0.05
	5 < Employees <10	0.751***	0.751***	0.751"**
		0.08	0.08	0.08
	Employees > 10	0.767**	0.767"*	0.767**
		0.09	0.09	0.09
	Outside debt ratio	0.812**	0.716***	0.715***
		0.08	0.09	0.09
	5,000 < Startup capital < 10,000	1.201**	1.035	1.033
		0.11	0.11	0.11
	10,000 < Startup capital < 25,000	1.007	0.944	0.941
		0.08	0.09	0.09
	25,000 < Startup capital < 100,000	1.111	0.999	0.994
		0.09	0.1	0.1
	Startup capital > 100,000	0.987	0.892	0.886
		0.1	0.1	0.1
	High credit risk	1.546***	1.547***	1.546***
		0.14	0.14	0.14
Innovation *measures*	High, medium tech	0.97	0.971	0.91G
		0.07	0.07	0.07
	Patents	1.299	1.297	1.318
		0.21	0.21	0.26
	Research&Development	0.272***	0.272***	0.297***
		0.03	0.03	0.04
			Interaction with 2008-2009 Recession	
Financing	5,000 < Startup capital < 10,000		1.619***	1.630***
			0.3	0.3
	10,000 < Startup capital < 25,000		1.264	1.278
			0.21	0.22
	25,000 < Startup capital < 100,000		1.430**	1.456**
			0.23	0.23
	Startup capital > 100,000		1.406**	1.436**
			0.24	0.25
	Outside debt ratio		1.423**	1.427**
			0.25	0.25
Innovation measures	Research&Development			0.741
				0.21
	Patents			0.949
				0.3
	N	4348	4348	4348
	arc	32637	32629	32632
	11	-16261.7	-16252.5	-16251

Note: Cox regression results are reported. The dependent variable is year of firm exit (or attrition) or whether the firm survives until 2011. Predictors are based on initial conditions in columns 1-3. *, **, *** denotes statistical significance at the 10, 5 and 1 to levels respectively. Other founder demographics, business characteristics, state and sector controls not shown.

Table 5: Probability of firm exit post-2009: effects of product innovation

	Exit post 2009, Ml	Exit post 2009, M2
Bachelor degree	-0.039*	-0.038*
	0.02	0.02
Graduate degree	0.028	0.029
	0.02	0.02
Number of past businesses	0.003	0.003
	0	0
Work experience	-0.021***	-0.021***
	0.01	0.01
0 < Employees < 5	-0.02	-0.021
	0.02	0.02
5 < Employees < 10	-0.027	-0.028
	0.03	0.03
Employees > 10	-0.027	-0.029
	0.04	0.04
Outside debt ratio	-0.068**	-0.064*
	0.03	0.03
5,000 < Startup capital < 10,000	-0.005	-0.008
	0.03	0.03
10,000 < Startup capital < 25,000	0	-0.002
	0.02	0.02
25,000 < Startup capital < 100,000	0	-0.002
	0.02	0.03
Startup capital > 100,000	0.036	0.034
	0.03	0.03
High credit risk	0.021	0.021
	0.04	0.04
High or medium tech	-0.036*	-0.038"
	0.02	0.02
Patents	0.096**	0.095**
	0.04	0.04
R&D	-0.076***	-0.085***
	0.02	0.02
Product innovation		
New to firm, 2009	-0.102***	
	0.02	
New to market, 2009		-0.092***
		0.03
N	2535	2535
aic	2489	2498
11	-1186.6	-1191.2

Note: Probit marginal effects reported. The dependent variable is a dummy equal to one if the firm fails in years 2010 or 2011, and zero otherwise. Estimation sample is conditional on firms that report surviving up until 2009; the first year survey questions on product innovation are introduced.
*, **, *** denotes statistical significance at the 10, 5 and 1 to levels respectively. Other founder demographics, business characteristics, state and sector controls not shown.

Table 6: Firm size and financing in startup year: explanatory factors

	Revenue			Employees			Finance		
Number of past businesses	-0.004	-0.004	-0.004	-0.008	-0.005	-0.006	0.017	0.019	0.018
	0.02	0.02	0.02	0.05	0.05	0.05	0.03	0.03	0.03
Years work experience	0.244***	0.243***	0.243***	0.301***	0.303***	0.305***	-0.169***	-0.168***	-0.168***
	0.04	0.04	0.04	0.08	0.08	0.08	0.06	0.06	0.06
0 < Employees ≤ 5	0.429***	0.431***	0.425***				0.531***	0.534***	0.540***
	0.10	0.10	0.10				0.16	0.16	0.16
5 < Employees ≤ 10	1.125***	1.131***	1.116***				0.995***	0.999***	1.013***
	0.17	0.17	0.17				0.29	0.29	0.29
Employees > 10	1.733***	1.739***	1.722***				1.820***	1.831***	1.847***
	0.22	0.22	0.22				0.34	0.34	0.34
outside debt ratio	-0.056	-0.058	-0.054	0.595	0.604	0.590			
	0.15	0.15	0.15	0.41	0.42	0.42			
5K < Startup capital ≤ 10K	0.330*	0.329*	0.330*	-0.312*	-0.309*	-0.308*			
	0.18	0.18	0.18	0.18	0.18	0.18			
10K < Startup capital ≤ 25K	0.466***	0.466***	0.463***	-0.238*	-0.231*	-0.216			
	0.14	0.14	0.14	0.14	0.14	0.14			
25K < Startup capital ≤ 100K	0.976***	0.975***	0.973***	0.210	0.220	0.234			
	0.14	0.13	0.13	0.25	0.25	0.25			
Startup capital > 100K	1.088***	1.088***	1.083***	2.099***	2.099***	2.143***			
	0.14	0.17	0.17	0.37	0.37	0.38			
High credit risk	-0.121	-0.123	-0.119	0.748**	0.764**	0.758**			
	0.24	0.24	0.24	0.34	0.34	0.34			
High medium tech	0.141	0.139	0.139	0.348	0.357	0.365	-0.478***	-0.471***	-0.471***
	0.11	0.11	0.11	0.24	0.24	0.25	0.17	0.17	0.17
Has patent	-1.074***	-1.071***	-1.089***	1.123	1.112	1.162	0.114	0.112	0.118
	0.35	0.35	0.35	0.71	0.71	0.71	0.50	0.50	0.50
R&D	0.021	0.030	-0.065	-0.181	-0.178	0.246	-0.091	-0.037	0.014
	0.13	0.13	0.21	0.36	0.37	0.66	0.20	0.20	0.30
Product innovation (2009-2011)									
New to firm	-0.062			0.429*			0.319**		
	0.10			0.25			0.15		
New to market		-0.111			0.507			0.165	
		0.12			0.35			0.19	
Interacted with R&D			0.105			-0.479			-0.013
			0.25			0.66			0.36
N	1784.000	1784.000	1784.000	1918.000	1918.000	1918.000	1991.000	1991.000	1991.000
R^2	0.39	0.39	0.39	0.15	0.15	0.15	0.20	0.19	0.19

Note: OLS regression results are reported. Dependent variable is log revenue of the firm in the first year of operation (columns 1-3); log number of employees in initial year (columns 4-6); log initial financing (columns 7-9). *, **, *** denotes statistical significance at the 10, 5 and 1 % levels respectively. Other founder demographics, business characteristics, state and sector controls not shown.

Table 7: Firm Growth: 2004-2011; Effects of product innovation

	Full period (2004 to 2011)			Pre (2004 to 2008)			Post (2008 to 2011)		
	Revenue	Employee growth (probit)	Employee growth (nbreg)	Revenue	Employee growth (probit)	Employee growth (nbreg)	Revenue	Employee growth (probit)	Employee growth (nbreg)
New to firm, 2009	0.650***	0.087***	0.462	0.009	0.124***	0.192	0.246***	0.073**	0.189
	0.23	0.03	0.32	0.29	0.03	0.3	0.08	0.03	0.25
New to firm, 2010	0.559*	0.122***	0.626	0.403	0.076*	-0.126	0.159	0.111***	0.342
	0.31	0.04	0.48	0.34	0.04	0.41	0.11	0.04	0.35
New to firm, 2011	0.569*	0.034	0.846	-0.327	0.100**	0.585	0.273**	0.069	0.454
	0.3	0.04	0.56	0.41	0.05	0.52	0.11	0.04	0.42
New to market, 2009	0.533*	0.127***	0.814**	-0.067	0.154***	0.352	0.224**	0.093***	0.519
	0.29	0.04	0.38	0.33	0.04	0.36	0.1	0.04	0.34
New to market, 2010	0.309	0.127**	0.465	0.301	0.128**	0.064	0.068	0.128**	0.239
	0.49	0.06	0.56	0.57	0.06	0.6	0.19	0.05	0.4
New to market, 2011	0.792*	-0.007	1.19	-0.869	0.11	0.692	0.334**	0.059	0.804
	0.46	0.06	0.88	0.7	0.07	0.83	0.16	0.06	0.75
Interacted with R&D, 2009	1.011*	0.220***	1.875***	0.437	0.222***	1.461**	0.338*	0.068	0.407
	0.54	0.06	0.64	0.5	0.06	0.64	0.18	0.06	0.63
Interacted with R&D, 2010	0.061	0.221**	2.191**	0.655	0.276***	1.444	-0.082	0.115	0.308
	0.84	0.09	1.01	0.52	0.09	1.01	0.31	0.09	0.85
Interacted with R&D, 2011	0.503	0.143	1.822	-0.073	0.201*	2.150*	0.231	0.026	0.088
	0.74	0.09	1.11	0.81	0.1	1.22	0.25	0.09	0.97
N	1770	1916	1916	1770	1916	1916	1770	1916	1916

Note: Columns 1, 4 and 7 report OLS estimates where the dependent variable is log annualized revenue growth. Columns 2, 5 and 8 report probit marginal effects where the dependent variable is whether the firm displays positive employment growth. Columns 3, 6 and 9 report negative binomial results where the dependent variable is the change in number of full-time employees. *, **, *** denotes statistical significance at the 10, 5 and 1 % levels respectively. Other founder demographics, business characteristics, state and sector controls not shown.

Table 8: Firm Growth: 2004-2011; Other effects

	Revenue	Employee growth (probit)	Employee growth (nbreg)
Number of past businesses	-0.017	-0.003	-0.033
	0.03	0.01	0.05
Years work experience	-0.068	-0.015	-0.11
	0.11	0.01	0.14
High, medium tech	-0.261	0.006	0.335
	0.23	0.03	0.34
0 < Employees ≤ 5	0.083	0.113***	-0.11
	0.26	0.03	0.3
5 < Employees ≤ 10	-0.027	0.247***	1.028*
	0.44	0.05	0.56
Employees > 10	0.423	0.189***	5.001***
	0.46	0.06	0.9
outside debt ratio	-0.334	-0.045	-0.243
	0.38	0.05	0.64
5K < Startup capital ≤ 10K	0.167	0.027	0.095
	0.35	0.05	0.34
10K < Startup capital ≤ 25K	0.134	0.133***	0.553**
	0.27	0.04	0.28
25K < Startup capital ≤ 100K	0.35	0.142***	0.559
	0.3	0.04	0.43
Startup capital > 100K	0.723**	0.109**	1.142**
	0.33	0.05	0.54
high credit risk	0.442	0.08	0.412
	0.65	0.08	0.75
Has patents	-0.072	-0.287***	-1.742*
	0.54	0.08	0.9
R$D	-1.036	0.02	-0.27
	0.75	0.09	0.84
Predicted hazard ratio	-0.933	-0.134	-1.579
	1.05	0.14	1.2
Constant	0.816		
	1.18		
N	1770	1916	1916
R^2	0.065		

Note: Growth measures based on full-period (2004 to 2011) as in previous table. Column 1 reports OLS estimates where the dependent variable is log annualized revenue growth. Column 2 reports probit marginal effects where the dependent variable is whether the firm exhibits positive employment growth. Column 3 reports negative binomial effects where the dependent variable is change in number of full-time employees.

Product innovation measures are not shown. *, **, *** denotes statistical significance at the 10, 5 and 1 % levels respectively.

Table 9: Annual outside debt raised, 2005 to 2011

	Surviving sample			Full sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Number of past businesses	0.035	0.035	0.032	0.012	0.012	0.011
	0.04	0.04	0.04	0.03	0.03	0.03
Work experience	0.006	0.01	0.017	0.016	0.021	0.025
	0.09	0.09	0.09	0.07	0.07	0.07
0 < Employees ≤ 5	0.572***	0.592***	0.608***	0.552***	0.554***	0.560***
	0.21	0.21	0.21	0.15	0.15	0.15
5 < Employees ≤ 10	1.185***	1.208***	1.222***	1.134***	1.142***	1.158***
	0.35	0.35	0.36	0.25	0.25	0.25
Employees > 10	1.720***	1.767***	1.808***	1.671***	1.693***	1.712***
	0.44	0.44	0.44	0.31	0.31	0.31
Outside debt ratio	1.563***	1.576***	1.592***	1.249***	1.252***	1.254***
	0.3	0.3	0.3	0.22	0.22	0.22
Outside debt	0.933***	0.931***	0.942***	0.818***	0.821***	0.825***
	0.21	0.21	0.21	0.14	0.14	0.14
5,000 < Startup capital ≤ 10,000	0.294	0.292	0.273	0.211	0.221	0.226
	0.34	0.33	0.34	0.21	0.21	0.21
10,000 < Startup capital ≤ 25,000	0.968***	0.999***	0.985***	0.818***	0.830***	0.832***
	0.24	0.24	0.24	0.15	0.15	0.15
25,000 < Startup capital ≤ 100,000	1.381***	1.407***	1.398***	1.329***	1.340***	1.344***
	0.28	0.28	0.28	0.18	0.18	0.18
Startup capital > 100,000	1.898***	1.915***	1.919***	1.919***	1.929***	1.941***
	0.3	0.31	0.3	0.2	0.2	0.2
High credit risk	-0.924*	-0.923*	-0.980*	-0.643*	-0.646*	-0.670*
	0.54	0.54	0.54	0.36	0.36	0.36
High or medium tech	-0.252	-0.226	-0.229	-0.310**	-0.301**	-0.298**
	0.2	0.2	0.2	0.13	0.13	0.13
Patents	-1.208*	-1.211*	-1.160*	-1.068***	-1.066***	-1.041***
	0.65	0.65	0.64	0.4	0.41	0.4
R&D	0.754	0.862	0.811	1.145***	1.257***	1.382***
	0.61	0.61	0.66	0.43	0.43	0.48
Product innovation						
New to firm, 2009	0.632***			0.785***		
	0.2			0.19		
New to firm, 2010	0.18			0.285		
	0.3			0.27		
New to firm, 2011	0.196			0.372		
	0.3			0.31		
New to market, 2009		0.418*			0.578**	
		0.24			0.23	
New to market, 2010		-0.325			-0.156	
		0.4			0.37	
New to market, 2011		-0.544			-0.381	
		0.44			0.43	
Interacted with R&D, 2009			0.673*			0.189
			0.39			0.34
Interacted with R&D, 2010			-0.628			-0.455
			0.65			0.59
Interacted with R&D, 2011			-0.25			-0.468
			0.66			0.64
Constant	4.093***	4.139***	4.047***	4.239***	4.339***	4.317***
	1	1	1.01	0.67	0.67	0.67
N	1916	1916	1916	4678	4678	4678
R^2	0.34	0.338	0.339	0.357	0.355	0.354

Note: OLS regression results are reported. The dependent variable is log average annual external debt (business loans) raised by the firm over the period 2005 to 2011 or until year of exit. *, **, *** denotes statistical significance at the 10, 5 and 1 % levels respectively. Full-sample regressions, columns 4-6, include year of exit dummies. Other founder demographics, business characteristics, state and sector controls not shown.

Table 10: Probability of applying for a business loan, 2007 to 2011

	Full sample			Surviving sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Number of past businesses	-0.003	-0.003	-0.003	-0.006	-0.006	-0.007
	0	0	0	0.01	0.01	0.01
Work experience	-0.003	-0.003	-0.003	-0.014	-0.013	-0.013
	0.01	0.01	0.01	0.01	0.01	0.01
0 < Employees ≤ 5	0.067***	0.067***	0.067***	0.063**	0.064**	0.066**
	0.02	0.02	0.02	0.03	0.03	0.03
5 < Employees ≤10	0.104***	0.103***	0.104***	0.123***	0.121***	0.124***
	0.03	0.03	0.03	0.05	0.05	0.05
Employees > 10	0.127***	0.125***	0.127***	0.062	0.06	0.068
	0.04	0.04	0.04	0.05	0.05	0.05
Outside debt ratio	0.066**	0.066**	0.066**	0.057	0.058	0.059
	0.03	0.03	0.03	0.04	0.04	0.04
Outside debt	0.011	0.011	0.013	0.027	0.026	0.029
	0.02	0.02	0.02	0.03	0.03	0.03
5,000 < Startup capital ≤ 10,000	0.018	0.019	0.022	0.032	0.034	0.034
	0.03	0.03	0.03	0.05	0.05	0.05
10,000 < Startup capital ≤ 25,000	0.027	0.027	0.029	0.053	0.055	0.055
	0.02	0.02	0.02	0.04	0.04	0.04
25,000 < Startup capital ≤ 100,000	0.086***	0.088***	0.090***	0.137***	0.141***	0.140***
	0.03	0.03	0.03	0.04	0.04	0.04
Startup capital > 100,000	0.112***	0.112***	0.115***	0.173***	0.174***	0.174***
	0.03	0.03	0.03	0.04	0.04	0.04
High credit risk	-0.005	-0.004	-0.004	0.018	0.02	0.016
	0.05	0.05	0.05	0.08	0.08	0.08
High or medium tech	-0.014	-0.013	-0.013	-0.006	-0.003	-0.005
	0.02	0.02	0.02	0.03	0.03	0.03
Patents	0.044	0.043	0.044	-0.008	-0.012	-0.004
	0.05	0.05	0.05	0.08	0.08	0.08
R&D	-0.002	-0.005	-0.013	-0.08	-0.083	-0.094
	0.06	0.06	0.07	0.09	0.09	0.1
New to firm, 2009	0.063***			0.066**		
	0.02			0.03		
New to firm, 2010	-0.011			-0.028		
	0.03			0.04		
New to firm, 2011	0.02			0.015		
	0.03			0.04		
New to market, 2009		0.084***			0.095***	
		0.03			0.03	
New to market, 2010		0.034			-0.013	
		0.04			0.05	
New to market, 2011		0.01			0.006	
		0.05			0.06	
Interacted with R&D, 2009			0.065			0.088
			0.04			0.06
Interacted with R&D, 2010			0.011			-0.049
			0.07			0.08
Interacted with R&D, 2011			-0.003			0.003
			0.07			0.09
Exit, 2008	-0.286***	-0.287***	-0.295***			
	0.03	0.03	0.03			
Exit, 2009	-0.152***	-0.153***	-0.161***			
	0.03	0.03	0.03			
Exit, 2010	-0.112***	-0.111***	-0.117***			
	0.03	0.03	0.03			
Exit, 2011	-0.062**	-0.062**	-0.064**			
	0.03	0.03	0.03			
N	3319	3319	3319	1916	1916	1916

Note: Probit marginal effects regression results are reported. The dependent variable is the probability a firm reports applying for a business loan from 2007 until 2011 or year of exit. Survey question on loan applications introduced in 2007; estimation sample is conditional on firms that survive up until 2007. *, **, *** denotes statistical significance at the 10, 5 and 1 % levels respectively. Full-sample regressions, columns 4-6, include year of exit dummies (relative to surviving firms). Other founder demographics, business characteristics, state and sector controls not shown.

Table 11: Probability of being denied for a business loan, 2007 to 2011

	Full sample			Surviving sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Number of past businesses	0.032***	0.030***	0.032***	0.026*	0.023	0.026*
	0.01	0.01	0.01	0.01	0.01	0.01
Work experience	-0.021	-0.02	-0.021	-0.008	-0.007	-0.01
	0.02	0.02	0.02	0.02	0.02	0.02
0 < Employees ≤ 5	-0.026	-0.023	-0.025	-0.034	-0.032	-0.039
	0.05	0.05	0.05	0.06	0.06	0.06
5 < Employees ≤10	0.06	0.058	0.054	0.044	0.038	0.027
	0.07	0.07	0.07	0.08	0.08	0.08
Employees > 10	0.069	0.072	0.077	0.009	0.01	0.013
	0.08	0.08	0.08	0.09	0.09	0.09
Outside debt ratio	-0.151**	-0.150**	-0.162**	-0.201**	-0.198**	-0.223***
	0.07	0.07	0.07	0.08	0.08	0.08
Outside debt	0.096*	0.092*	0.106**	0.086	0.082	0.101*
	0.05	0.05	0.05	0.06	0.06	0.06
5,000 < Startup capital ≤ 10,000	0.139	0.133	0.135	0.063	0.049	0.058
	0.09	0.09	0.09	0.1	0.1	0.1
10,000 < Startup capital ≤ 25,000	-0.048	-0.043	-0.046	-0.157*	-0.150*	-0.148*
	0.07	0.07	0.07	0.08	0.08	0.08
25,000 < Startup capital ≤ 100,000	0.044	0.045	0.041	-0.009	-0.007	-0.01
	0.07	0.07	0.07	0.08	0.08	0.08
Startup capital > 100,000	-0.097	-0.092	-0.103	-0.095	-0.09	-0.1
	0.08	0.08	0.08	0.09	0.09	0.09
High credit risk	0.185*	0.183*	0.183*	0.175	0.184	0.164
	0.1	0.1	0.1	0.12	0.12	0.12
High or medium tech	-0.05	-0.04	-0.051	-0.017	0.001	-0.013
	0.05	0.05	0.05	0.06	0.06	0.06
Patents	0.292***	0.302***	0.289***	0.372***	0.375***	0.355***
	0.11	0.11	0.11	0.12	0.13	0.13
R&D	0.094	0.105*	0.032	0.133**	0.148**	0.016
	0.06	0.06	0.09	0.07	0.07	0.1
New to firm, 2009	0.088			0.114*		
	0.06			0.06		
New to firm, 2010	0.153*			0.199**		
	0.08			0.08		
New to firm, 2011	0.003			0.055		
	0.09			0.09		
New to market, 2009		0.086			0.111	
		0.06			0.07	
New to market, 2010		0.101			0.181*	
		0.1			0.1	
New to market, 2011		-0.077			-0.038	
		0.12			0.12	
Interacted with R&D, 2009			0.123			0.191*
			0.1			0.11
Interacted with R&D, 2010			0.279*			0.483***
			0.16			0.15
Interacted with R&D, 2011			0.188			0.264
			0.19			0.19
Exit, 2008	0.051	0.035	0.022			
	0.12	0.12	0.12			
Exit, 2009	-0.031	-0.045	-0.043			
	0.09	0.09	0.09			
Exit, 2010	-0.142	-0.152*	-0.159*			
	0.09	0.09	0.09			
Exit, 2011	-0.015	-0.023	-0.015			
	0.07	0.07	0.07			
N	3319	3319	3319	1916	1916	1916

Note: Heckman selection probit marginal effects regression results are reported. The dependent variable is the probability a firm reports being denied a business loan at least on one occasion from 2007 until 2011 or year of exit. Survey question on loan applications introduced in 2007; estimation sample is conditional on firms that survive up until 2007. *, **, *** denotes statistical significance at the 10, 5 and 1 % levels respectively. Full-sample regressions, columns 4-6, include year of exit dummies (relative to surviving firms). Other founder demographics, business characteristics, state and sector controls not shown.

7 A Appendix

Table A.1: Technology Sampling Strata Definitions

Technology Sampling Stratum	SIC Code	Industry
High	28	Chemicals and allied products
	35	Industrial machinery and equipment
	36	Electrical and electronic equipment
	38	Instruments and related products
Medium	131	Crude petroleum and natural gas operations
	211	Cigarettes
	229	Miscellaneous textile goods
	261	Pulp mills
	267	Miscellaneous converted paper products
	291	Petroleum refining
	299	Miscellaneous petroleum and coal products
	335	Nonferrous rolling and drawing
	348	Ordnance and accessories, not elsewhere classified
	371	Motor vehicles and equipment
	372	Aircraft and parts
	376	Guided missiles, space vehicles, parts
	379	Miscellaneous transportation equipment
	737	Computer and data processing services
	871	Engineering and architectural services
	873	Research and testing services
	874	Management and public relations
	899	Services, not elsewhere classified
None		All other

Source: KFS Baseline Methodology Report.

Table A.2: High Tech vs low tech

	Non-Tech	Medium, High-Tech
Bachelor	0.24	0.28
Graduate	0.20	0.37
Years Work Experience	10.1	13.9
Startup FT Employees	2.95	3.30
Revenue > 0	0.64	0.66
Revenue ($)	228,495	318,927
ROA	0.52	0.66
Total Assets ($)	372,639	216,289
Patents	0.02	0.05
R&D	0.16	0.27
New Product	0.25	0.31
Univ Coop	0.09	0.12
Survive to 2011	0.44	0.53
N	2894	2034

Note: Based on 4,298 firms in baseline year. Numbers reported are shares or averages dollar amounts. Statistics are all significantly different at 5 percent level across columns.

Table A.3: Firm Exits

	Unweighted		Weighted	
	Survive	Fail	Survive	Fail
2005		301		4653
2006		366		5979
2007		343		5330
2008		383		5752
2009		293		4443
2010		249		3807
2011	2007	274	28109	4379
Total	2007	2209	28109	34343

Note: Based on 4,216 firms in baseline year in Kauffman Firm Survey.

Table A.4: Composition of Startup Firms by Industry

		Startup Year		**Survive Until 2011**	
Industry	NAICS	Unweighted	Weighted	Unweighted	Weighted
Construction	23	308	6503	134	2817
Manufacturing	31, 32, 33	589	3734	304	1758
Wholesale Trade	42	186	3426	85	1531
Retail	44, 45	465	9332	165	3263
Transportation, Warehousing	48,49	96	1804	41	761
Information	51	142	1967	66	844
Finance, Insurance	52	164	3376	69	1411
Real Estate, Rental, Leasing	53	159	3310	84	1720
Professional Services	54,55,61	1091	10943	590	5894
Waste Management, Remediation	56	306	6080	141	2820
Health Care, Social Assistance	62	95	1937	41	798
Arts, Entertainment, Recreation	71	92	1536	47	724
Accomodation, Food Services	72	88	1740	31	590
Other Services	81, 11, 21 22, 92	435	6766	209	3178
Total		4216	62452	2007	28109

Note: Based on 4,216 firms in baseline year in Kauffman Firm Survey.

Table A.5: Capital Injections: Surviving vs Exiting Firms

		All Firms	Mean if > 0	Count	Survive	Fail
2005	Owner Equity	10,157	29,358	1878	9,912	9,337
	Owner Debt	2,310	11,899	933	2,096	2,679
	Outside Debt	22,503	83,239	1318	23,733	21,602
2006	Owner Equity	5,584	25,562	1283	6,720	3,909
	Owner Debt	1,926	11,953	812	2,067	1,837
	Outside Debt	25,063	81,709	1330	25,283	19,948
2007	Owner Equity	3,587	24,766	941	4,970	1,639
	Owner Debt	1,513	13,547	631	2,106	1,138
	Outside Debt	20,383	82,280	1194	23,647	12,037
2008	Owner Equity	3,056	27,700	770	5,321	866
	Owner Debt	1,420	14,212	575	2,333	636
	Outside Debt	19,015	95,296	1077	27,105	6,155
2009	Owner Equity	1,902	27,458	590	4,175	271
	Owner Debt	1,082	13,453	505	1,868	143
	Outside Debt	14,733	94,466	932	25,904	2,067
2010	Owner Equity	1,418	22,021	491	3,073	10
	Owner Debt	750	14,781	386	1,542	4
	Outside Debt	9,845	99,554	756	22,081	207

Note: Based on sample of 4,216 firms in Kauffman Firm Survey. Numbers displayed are in average dollar terms.

Firm R&D and Knowledge Spillovers*

Kim P. Huynh† Zeno Rotondi‡

Leonard Sabetti§

March 20, 2020

Abstract

Private investment in research and development (R&D) is risky and sub-optimal in aggregate; public sector driven R&D, such as those made in the university sector, can play a role in crowding in private sector investment through knowledge spillovers. Using data on a large number of Italian manufacturing firms over the period 1995 to 2003, we find evidence that geographical variation in the R&D intensity of higher education sector, a good proxy for the local intensity of knowledge spillovers, affects R&D spending decisions of firms, particularly for larger firms and firms in the high-tech sector. Further, crowding-in effects from firms cooperating with universities directly on R&D are estimated at roughly 30 percentage points. In contrast and in line with previous studies, we do not find any effects from local banking development or bank finance; compared to the importance of internal funds and government subsidies as main sources of R&D financing. In light of our findings, expanding the presence of R&D intensive universities or promoting a better spatial match between university R&D and local industry could lead to both direct and indirect impacts towards boosting measures of private sector R&D intensity.

Key Words: R&D, University-Industry Cooperation, Knowledge Spillovers, Geography.

JEL Classification: D21, D24, G21, G28, O31, O32, O38.

*Previous version of this paper circulated as Huynh, Kim P. and Rotondi, Zeno, R&D Spending and Knowledge Spillovers (July 15, 2009). Available at SSRN: https://ssrn.com/abstract=968666.

We thank Michele Fratianni, Luigi Benfratello, Patrizio Bianchi, Luigi Guiso, Marco Pagano, Fabio Schiantarelli and seminar participants of XV International Tor Vergata Conference on Money, Finance and Growth for helpful comments and suggestions. We are grateful to Luigi Guiso for providing part of the data used in this study. We are also grateful to Tony Riti for helpful information on the Capitalia Survey. The usual disclaimer applies.

† Bank of Canada. E-mail: kim@huynh.tv.

‡ UniCredit S.p.A., Via Zamboni, 20, Bologna, Italy 40126. Email: zeno.rotondi@unicreditgroup.eu.

§CERDI, Universite´ Clermont Auvergne (France) Email: leonard.sabetti@etu.udamail.fr

1 Introduction

Theories of economic growth have largely emphasized the importance of R&D for enhancing productivity and creating well-paying jobs, especially for industrialized countries which tend to stagnate after picking all the low-hanging fruit. Given the positive externalities, aggregate wide spending on R&D is likely to be below the socially optimal level in many countries. For example, Bloom, Schankerman, and Reenen (2005) find that the public benefits of firms engaging in R&D were roughly double the private benefits using a sample of U.S. firms. While public sector remains an important source of R&D financing through transfers and tax cred- its, tightening budgets suggests that policymakers may need to seek alternative measures to promote R&D spending among firms.[1] For example, the Lisbon Agenda in Europe called for achieving a target ratio of R&D expenditures to GDP of three percent, with two thirds financed by the business sector.

At the micro level, firm R&D is considered an important input into the innovation process whereby firms introduce new products or processes; see as an example Mairesse and Mohnen (2005). Understanding determinants of R&D spending, which vary across the stage of a firm's life- cycle and particularly as its financial position changes, remains important for understanding firm performance. For instance, Phillips and Zhdanov (2012) show how large firms are more likely to acquire small innovative firms than to invest in R&D themselves in an active acquisition market. Rosenberg (1990) studies what types of American firms invest in basic research with their own money and shows that it is highly concentrated in a select number of industrial sectors and dominated by a small number of typically large firms within those sectors.

One important channel for promoting firm R&D expenditures remains the influence of university research as well as university-industry cooperation, as documented in Feller (1990), Adams, Chiang, and Starkey (2001), Hall, Link, and Scott (2003), and Siegel and Phan (2004). As discussed in Mairesse and Mohnen (2010), cooperation in R&D helps firms share knowledge, benefit from complementarities and to reduce risks or save costs. The geography of university-industry cooperation may also play an important role given evidence that knowledge spillovers are mostly localized, for example Jaffe, Trajtenberg, and Henderson (1993), Adams (2002). Another important dimension is the ability of the private sector to absorb and apply the ideas arising out of the academic spillovers from the higher education sector. For example, Agrawal

and Cockburn (2002) show that the presence of a large, local, R&D-intensive firm makes local university research more likely to be absorbed and stimulates local industrial R&D (anchor tenant hypothesis). In light of these findings, understanding the linkages between the location of resources and the effectiveness of R&D, as well as the role of university-industry cooperation, could be important for addressing stagnant productivity. More recently, Arora, Belenzon, Patacconi, and Suh (2019) document the changing structure of American innovation whereby the decline in the corporate R&D lab is being replaced by university R&D as a principle source for new ideas.

We shed light on these issues and contribute to the literature using firm-level survey data on Italian manufacturing firms where R&D activities are reported including whether jointly undertaken with universities. The data allow us to test for the crowding-in hypothesis as well as the role of broader spillover effects. We estimate the impact from university cooperation on the intensive margin of firm R&D expenditures using propensity score matching techniques under the selection on observables assumption. We find an effect of roughly 5 percentage points and roughly equivalent to one third of the effect from public subsidies.

We also uncover a role from underlying regional knowledge spillovers highlighting the role of indirect effects. Because we find that variation in R&D intensity of the higher education sector across regions influences the probability of firm-university linkages, local measures of R&D intensity of the higher education sector (HES) serve as a suitable proxy for local knowledge spillovers. Our regressions reveal that the local intensity of university-knowledge spillovers positively affects firm R&D expenditures and are robust across sub-samples (including when considering the sub-sample of firms that have no direct links to universities) and using instrumental variables. In contrast, we do not find any positive effect from the local supply of banking or credit. Our results suggest the presence of a spatial mismatch in Italy between R&D intensive firms located disproportionately in the north and R&D intensive universities located primarily near the government center in Rome. The richness of the data, including the information on sources of R&D financing, underly our contribution to the literature.

While the focus of our paper is on the extensive and intensive margins of firm R&D expenditures, previous studies have used the same dataset to study determinants of firm innovation. For example, Benfratello, Schiantarelli, and Sembenelli (2008) find that local banking development plays an important role in fostering firms' innovative activities, but the effect is stronger

for process innovation than product innovation. Herrera and Minetti (2007) find that informed banks, measured by the duration of credit relationships, foster innovation by financing the introduction of new technologies while in contrast they find no effect on the financing of internal research.

The remainder of the paper is organized into the following sections. Section 2 provides some background and historical context related to policy issues in the US and Europe and describes how the institutional environment of Italy provides a good laboratory to address our research questions. This section also contains an examination of the available data on R&D for Italy in comparison with other countries and across Italian regions. Section 3 describes our empirical methodology and section 4 provides an analysis of the data used and the measurement of the variables. Section 5 presents estimation results and robustness checks. Section 6 concludes.

2 Background and institutional setting

The debate on university-industry cooperation beginning in the early 1980s has led to a number of contributions which have influenced policy in the United States. For example, the Bayh-Doyle Act of 1980 established the right for universities to patent inventions resulting from federally funded research. The Economic Recovery Tax Act of 1981 granted firms an R&D tax credit for company-financed academic research.[13] Finally, the Small Business Innovation Research Act of 1982 (SBIR) promoted agency-financed start-ups, including those headed by university researchers.

Despite the ambitions of the Lisbon Agenda, structural reforms among many European universities remain absent, for instance see Jacobs and van der Ploeg (2006). Specifically, Bianchi and Ramacciotti (2005) describe the long period of incomplete reforms among the Italian higher education sector. On one hand, reforms ensuring university autonomy in managing its own resources and organization were introduced in the 1990s. On the other hand, the normative framework on intellectual property rights attributes the property rights exclusively to the university researcher, weakening the mechanism that allows universities to finance their

[13] However, many small startup firms which typically do not earn profits in first few years of operation and whose survival rate is uncertain arguably cannot readily claim this tax credit.

development by means of the industrial spillovers of their research.

The focus of this paper is the period 1995-2003, which corresponds to when our data was collected. Figure 1 reports the R&D intensity for developed countries in 2003. Italy's R&D intensity is roughly one percentage point below the EU25 average (two percent) and roughly 1.5 percent below the R&D intensity of the United States. While the Lisbon Agenda called for a target of two-thirds financing of R&D by the business sector, Italy's ratio is roughly 50 percent, slightly below the EU25 average.

Figure 2 depicts the significant variation in R&D intensity across Italian regions. Lazio and Piemonte have R&D intensity ratios roughly at the two percent level (in line with EU25 aver- age), followed by seven regions with values close to one percent and the remaining half of the regions fall well below one percent. Further, there has been only modest improvement over the period since 1995. Interestingly, several regions with the most developed local economies, such as Piemonte, Lombardia and Emilia Romagna, are characterized by R&D intensity levels close to one percent, in line with less developed local economies such as Campagna. In contrast, Lazio is usually included among the less developed regions (see Guiso, Sapienza, and Zingales (2004a)), but it has one of the highest levels of R&D intensity. The right-hand side plot shows the R&D intensity of the higher education sector is lower in regions with more developed local economies, such as Piemonte, Lombardia, Veneto, and Emilia Romagna (implying a greater of the business sector). This data suggests that where firms would benefit from local university knowledge spillovers, there are fewer financial resources for the local higher education sector. While not shown, more developed but less R&D intensive regions are also those with the *higher* measures of patent intensity. They have experienced increases in the rate of patent in- tensity of 50 percent or greater over the period 1995 to 2003, and yet R&D expenditures have remained mostly unchanged. This may not be surprising in light of Cohen, Nelson, and Walsh (2000) who find that U.S. manufacturing firms tend to protect profits from their inventions more so by secrecy and lead time as opposed to patents, although less so far large firms who tend to rely on patents as a form of barrier to entry. Tabarrok (2002) describes the divergence between the actual effects of patent system and its supposed intended effects of incentivizing firms to engage in R&D by allowing them to recoup sunk costs. The increasing patent intensity in the north may be crowding out R&D spending in more R&D intensive regions.

In the remainder of the paper, we conduct an empirical analysis of the causal links between university-firm direct cooperation and local university knowledge spillovers on firm R&D intensity.

3 Research Design

The firm's intensive margin decision of R&D expenditure can be modeled as:

$$y_i = \alpha_i x_i + z_i \delta_{11} + u_i, \tag{1}$$

where y_i is the amount of R&D expenditure of firm i, x_i is a vector of firm-level control variables, z_i consists of potentially endogenous variables and u_i is a stochastic error term. Our set of endogenous variables of interest are in the form of both direct and indirect factors that could influence firm R&D investment decisions. Uncovering the causal nature of these variables and their estimated magnitudes are important for policy implications.

3.1 Direct effects: firm-university cooperation

We couch the role of university-industry cooperation in enhancing R&D expenditures in terms of a treatment effect according to the program evaluation approach; see Rosenbaum (2002) for an overview. While firms are not randomly assigned to collaborate with universities on R&D, the potential causal impact of such treatment can be uncovered via comparing treated firms to similar non-treated firms using propensity score matching. The approach is similar to estimating a two-system equation of the outcome and selection mechanism. In both cases, the use of exclusion restrictions help provide identification. Panel data methods are an alternative way to account for unobserved heterogeneity.

The ATE ($\tau_{ATE|X}$) represents the expected treatment on a randomly drawn individual across the entire sample whereas the ATT ($\tau_{ATT\ |\ X}$) measures the mean effect for the sample in the treatment group. From a policy perspective, the latter quantity may be more relevant because it excludes individuals in the sample who are not very likely to become part of the treatment group. In order to estimate $\tau_{ATE|X}$ and $\tau_{ATT\ |X}$ we rely on the propensity score, $p(I_i = 1|\mathbf{X_i})$, which denotes the probability that individual i entered the treatment group conditional on observable characteristics. This technique, known as propensity-score matching (PSM), simplifies the

estimation of the desired causal impact by reducing the matching problem from a vector of observable characteristics to a simple unit-dimensional measure. Firms in the treatment group are matched to firms in the control group by the estimated propensity score, $\hat{p}(X_i)$, based on some distance metric. In our implementation, we use both nearest-neighbour matching (NN) and kernel matching (KM). The PSM estimators for $\tau_{ATE|X}$ and $\tau_{ATT\ |X}$ are defined as,

$$\widehat{ATE} = \hat{\tau}_{ATE|X} = \frac{1}{N}\sum_{i=1}^{N}\frac{[I_i - \hat{p}(X_i)]y_i(I_i)}{\hat{p}(X_i)[1-\hat{p}(X_i)]}, \tag{2}$$

$$\widehat{ATT} = \hat{\tau}_{ATT|X} = \frac{1}{N_1}\sum_{i=1}^{N_1}\frac{[I_i - \hat{p}(X_i)]y_i(I_i)}{1-\hat{p}(X_i)}, \tag{3}$$

where N_1 is the number of firms in the treatment group while N_2 are the number of firms in the control group; N defines the total number of firms in the sample. These estimators can be interpreted as an average of firm R&D expenditures weighted by propensities to cooperate with universities on R&D.

Overall, inference hinges on the quality of the matching, known as the balanced covariate or overlap condition, as well as the lack of selection on unobservables. As discussed in Mairesse and Mohnen (2010), a large number of studies on the determinants of cooperation on innovation related activities finds that firm size, technology sector, incoming spillovers and appropriability are some of the key explanatory factors; see for example Veugelers and Cassiman (2005).

3.2 Indirect effects: regional knowledge spillovers and local supply of credit

Instrumental variables (IV) are used to account for endogeneity in estimating the relation between firm R&D expenditures and local factors categorizing the firm's environment. Using two-stage least squares approach, w_i are a set of variables that explain the endogenous variable z_i, but are uncorrelated with the error term u_i in equation (1). The effect of these instruments z_i is captured by parameters δ_{21} in the following equation,

$$z_i = w_i\delta_{21} + v_i. \tag{4}$$

To ensure the validity of our instruments, we also perform diagnostic checks. We test the as-

sumption that our instrument is uncorrelated with the true error term in the first equation using the Hansen-Sargan over-identification restriction. Rejecting this test of over-identification can be interpreted as a failure of the instruments to satisfy the orthogonality conditions.

In this section we are interested in studying the role of regional R&D intensity of the higher-education sector in Italy as a treatment variable (indirect effect), as we find that it serves as a good proxy for the local intensity of knowledge spillovers. Specifically, we show that there is a significant regional effect due to local differences in the financing of basic research on the probability that a firm cooperates with a university on R&D activities. As a result, regional differences are also likely to influence firm investment decisions - even when no direct link to a university exists - via the potential for knowledge spillovers. For instance, this channel could work via including a local firm to invest in R&D to absorb and acquire knowledge in order to stay ahead of its competitors. For robustness, we instrument this treatment variable using a set of variables that reflect features of the regional structure of higher education: the per capita total number of universities and the per capita number of universities with faculty of science in the region in 1995. These instrumental variables should account for variation in the treatment variable but should be orthogonal to our outcome variable.

As an additional indirect factor, we consider the role of local supply of credit as a potentially important determinant of firm R&D spending. If firms in certain regions are more likely to obtain financing on more generous terms, then they might also be more likely to engage in greater R&D spending. Without controlling for this effect, our results could be biased. In Italy, banks are the main source of external finance with a smaller role played by private equity firms relative to in the United States where typically credit markets are more advanced. An advantage of focusing on evidence from Italy is the delimitation of the banking system by local regions, as pointed out in Guiso, Sapienza, and Zingales (2004a). As a result, we are able to make use of exogenous restrictions on local supply of finance.

In particular, following Guiso, Sapienza, and Zingales (2004a), banking development is instrumented with a set of variables that describe the banking market as of 1936. The indicator of banking development, a measure of households' likeliness of being credit rationed, is constructed based on data from the Survey of Households Income and Wealth (SHIW) over the 1990s. The survey is conducted by the Bank of Italy and is one of the few comprehensive surveys that asks households whether they have been denied credit or have been discouraged

from applying and additionally contains precise information on respondents' location. The authors show that access to credit in the 1990's can be explained by the level and composition of supply of credit in 1936. The year 1936 coincides with the introduction of a banking law that subsequently greatly influenced the availability of branches 50 years later. The constructed indicator of banking development is correlated with those variables that describe the nature of the banking market in 1936: the share of bank branches owned by local banks, the number of saving banks in the region, the number of cooperative banks in the region and the bank branches in the region.

We estimate regressions on the pooled firm level data covering the 1995-2003 period. Since our main explanatory variables of interest, banking development and R&D intensity of HES, vary only at the regional level, standard errors are adjusted for the possible dependence of the residuals within regional clusters.

4 Data

We use the "VIII Indagine sulle Imprese Manufatturiere", a survey of manufacturing firms conducted by the Italian banking group Capitalia-Mediocredito Centrale. Our analysis builds on three waves run in 1998 (covering the period 1995-1997), 2002 (covering the 1998-2000 period) and 2004 (covering the 2001-2003 period). The resulting samples are stratified by firm size (number of employees), by sectors (four sectors according to Pavvit taxonomy) and by geographical area (North and Center-South). The samples are representative of Italian manufacturing firms with more than 10 employees and contain the universe of firms with over 500 employees. Each sample comprises over 4000 firms. As our primary interest, the survey investigates whether firms engage in R&D and whether they cooperate with universities on R&D expenditures. It also contains details on firm financing conditions, owner demographics and other variables. In general, the survey is highly reliable and comprehensive as it is used for providing information for the strategies of the banking group as well as for public policies promoting firm competitiveness, see Minetti and Zhu (2011) for more details.

Table A.1 in the appendix provides descriptive statistics of the variables used from the survey. Firms' R&D intensities are R&D expenditure ratios relative to production. R&D expenditure is deflated with a weighted average of the hourly earnings in manufacturing index

and the aggregate business investment price index, where the weights used are respectively 0.9 and 0.1, as in Parisi, Schiantarelli, and Sembenelli (2006). Production is computed as the sum of sales, capitalized costs and the change in work-in-progress and in finished goods inventories, with all variables deflated with the appropriate production price index. South is a dummy for regions south of Rome. Firm's age is the log of the number of years from its birth. Firm size is the log of number of employees. We have considered also a set of variables describing the sources of finance for investment in R&D: internal funds, new equity, fiscal incentives, public transfers, bank debt (subsidized or not).

Following Guiso, Sapienza, and Zingales (2004a), the maximum rate of growth internally financed is given by $maxg = ROA/(1 - ROA)$, where ROA is return on assets. Variables like ROA, sales, capitalized costs, and the change in work-in-progress and in finished goods inventories, which were derived from firms' balance sheets, are from AIDA databank.

Table 1 provides an overview of firm adoption rates of R&D in our sample, broken down by firm size, age, as well as whether it is in a high-tech sector. Roughly 35 percent of firms invest in R&D, whereas adoption rates tend to increase with firm size, age and in the high-tech sector. Further, over time, we see a general increase in R&D adoption from 30 percent to 40 percent over the three waves. When we condition on the share of firms engaging in R&D, we observe that roughly 12 percent have ties or cooperate with universities; which remains fairly constant over the time period under consideration. In contrast, the share of R&D firms receiving fiscal incentives grows from 5 percent to 10 percent; while the share of firms receiving public transfers increases from 10 percent to roughly 19 percent. As a result, one hypothesis for the growth of R&D adoption is driven by public-sector financial incentives. Table 1 also reports average firm R&D intensity rates across the three waves, broken down by whether a firm cooperates with a university or receives financial incentives. For the sample of benchmark firms which do not receive any external support, the average R&D intensity measure is lowest at roughly 1.7 percent. Firms that report cooperating with a university on R&D activities on average display R&D intensity rates at 2.5 percent. Finally, firms with the highest intensity rates not surprisingly, roughly 3.5 percent on average, receive financial support in some manner. The data suggest an important role for financial incentives but also for universities in influencing firm R&D expenditures.

Table 2 reveals the main source of financing for R&D investment is internal funds, with 80 percent on average across the three waves. In contrast, bank lending is considerably less important, suggesting that banking development may play a minor role compared to that for fixed investments (see Benfratello, Schiantarelli, and Sembenelli (2008)). This can also be seen from the relatively large share of firms with 100 percent internal finance and with no bank finance, roughly 66 percent and 86 percent respectively on average across the three waves. While on average the share of R&D expenditures financed by public transfers or fiscal incentives remain low, under ten percent, the share financed is much higher for the subset of firms actually receiving either type of financial incentive.

Finally, we augment the survey with regional data from the Italian Bureau of National Statistics (ISTAT) and from Guiso, Sapienza, and Zingales (2004a), Guiso, Sapienza, and Zingales (2004b). From ISTAT, we take the R&D intensity of higher education sector (HES) 1995, the per capita total number of universities and the per capita number of universities with faculty of science in the region in 1995, and the GDP per capita in Euro prices in 1991. Per capita GDP is the log of per capita net disposable income in the province in 1991.

From the dataset used in Guiso, Sapienza, and Zingales (2004a) and Guiso, Sapienza, and Zingales (2004b), we use the following regional variables: share of bank branches owned by local banks, number of savings banks per 10,000 inhabitants, number of cooperative banks per 10,000 inhabitants and bank branches per 10,000 inhabitants in the region. These variables describe the banking market in Italy as of 1936. We use the same indicator of local banking development computed by Guiso-Sapienza-Zingales. They show that the determinants of the geographical differences in the degree of financial development are those variables that describe the structure of the banking market in 1936. Social capital is measured by average voter turnout at the provincial level for all referenda in the period between 1946 and 1987. Judicial efficiency is the number of years it takes to have a first-degree judgement in the province.

5 Results

5.1 University-firm R&D cooperation

Our first set of results model the probability a firm cooperates with a university on R&D. Using the propensity score and based on matching techniques, we will then obtain measures of a treatment effect on the firm's rate of R&D expenditure that accounts for selection. As a by-product of the first-stage estimation, we also test whether the regional R&D intensity of HES has any predictive power in explaining local cooperation between universities and firms; providing evidence on whether such a measure could serve as a good proxy for knowledge spillovers. Table 3 presents our estimates of the probability that a firm cooperates with a university on R&D activities. Column 1 reports the Probit estimates of the impacts of firm specific and regional characteristics on the probability of cooperation with university on R&D activities. The regional R&D intensity of HES is statistically significant and has a positive impact. The marginal effect of a one percent increase leads to a 0.09 increase in the probability of university-industry cooperation. The firm specific characteristics are significant with the expected signs, such as firm size and age, while social capital and per capita GDP are not significant. Column 2 reports the estimates obtained using the total number of universities and the number of universities with faculty of science by region in 1995 as instruments. Column 3 applies an instrumental variable approach along with a probit model to provide better model specification.

The results confirm the importance of local differences in the financing of basic research in affecting the probability of cooperation with university on R&D activities (at the five percent level). Furthermore, this result offers some support to our indirect measure of university knowledge spillovers. Moreover, the Hansen-Sargan test does not reject the joint null hypothesis that the instruments are valid. The IV estimation method relies on the assumption of a linear probability model for university-industry cooperation. Therefore, in column 3, we also provide the estimates derived from a conditional maximum likelihood (IV-Probit) technique proposed in Wooldridge (2008) which does not require the assumption of a linear probability model. This technique uses maximum likelihood to estimate a probit model in the presence of an endogenous variable. The Wald test of exogeneity indicates that we cannot reject the null

hypothesis of no correlation between the first stage and second stage regression. This result provides some evidence that the regional R&D intensity of HES is exogenous. These robustness checks confirm the importance of local differences in the financing of basic research in affecting the probability of firms' cooperation with university on R&D activities (still at the five percent level). The results are similar to what is found by Veugelers and Cassiman (2005).

5.2 Determinants of Firms' R&D Expenditures

We study the determinants of firm R&D expenditures based on the dependent variable in equation 4; measured in terms of the deflated logarithm of R&D expenditures which are greater than zero.

5.2.1 Direct effects from university cooperation

Using the propensity score estimates derived previously, we can obtain estimates for the treatment effect on firm R&D expenditures on average (ATE) and conditional on the sample of treated firms (ATT). Table 4 presents the results including estimates from a simple linear model (OLS). Overall, our findings point to an effect of a roughly 30 to 40 percentage point increase.

As a back-of-the-envelope estimate, we observe that the average effect for the treated firm is roughly equal in magnitude to the average amount of R&D dollars financed by government subsidies (public transfers or fiscal subsidies). While we lack granular information about the degree of cooperation between the firm and university, the economic effects should be of interest to policymakers given that crowding in occurs without financial incentives. Further, it would be interesting to study the impacts of such induced R&D on firm outcomes relative to a counter- factual; such as measures of firm growth or productivity.

5.2.2 Indirect effects from local factors

We examine differences in the intensive margin of investments in R&D induced by local differences in banking development and R&D intensity of HES, controlling for firm specific observables. Besides calendar year and industry dummies, as control variables, we use a combination of individual, provincial and regional characteristics. For individual characteristics of the firm, we use measures of internally financed maximum growth, fiscal subsidies (in order to con-

trol for sources of finance alternative to bank debt) and firm size; among other variables. For provincial characteristics we have per capita GDP and the level of social capital.

Table 5 presents the results. Column 1 reports the OLS regression estimate of the impact of banking development on the amount of R&D expenditures. The indicator of banking development is not statistically significant. The individual characteristics are all statistically significant and have the expected effect while the provincial and regional characteristics are not significant. Column 3 estimates the same specification for the control variables but now examines the impact of R&D intensity of HES on the amount of R&D expenditures. Opposite to banking development, the regional R&D intensity of HES is highly significant (at the 1 percent level) and has a positive effect on the amount of R&D investments. This finding might reflect an important role played by the local presence of universities through local knowledge spillovers.[3] From column 3 we can observe that again the individual characteristics are all statistically significant and have the expected effect. While among the significant provincial and regional characteristics, we include per capita GDP and an indicator for the South.

In order to account for potential endogeneity problems, we estimate the model with IV, the specification used in columns 1 and 3. In column 2, we can see that the estimated coefficients are almost the same of those reported in column 1 and that the coefficient of the indicator of banking development is still not significant. Similarly, we can see from column 4 that the IV estimation confirms the significant impact of R&D intensity of HES on the amount of R&D expenditures. The mean of local R&D intensity of HES is 0.38 percent of GDP. Focusing on the results of column 4, a quantitative magnitude can be calculated for the impact of an increase in the local R&D intensity of HES. The impact of increasing R&D intensity of HES by 1 per- cent will result in an increase of about 0.97 percent increase in R&D expenditures for the firm. Similar conclusions are reached in column 5, where we estimate by means of IV a specification which accounts for both local factors under examination. In all cases, as it is possible to see from columns 2, 4 and 5 the Hansen-Sargan test cannot reject the joint null hypothesis that the instruments are valid.

The rationale for the finding of the irrelevance of banking development is rooted in the high

[3]The result is robust to excluding the sample of firms that report cooperating with universities on R&D with a coefficient of 0.90 and same degree of statistical significance.

degree of risk and the complexity of evaluating future prospects of activities related to innovation. This uncertainty makes banking intermediaries not ideal for financing R&D expenditures. This increases the probability of firms, especially high-tech ones, of being credit constrained with respect to investments in non-tangibles (see Guiso (1998)). Additionally, as shown in the present analysis, this is true independently of the degree of credit market development.

This argument is consistent with the findings of Herrera and Minetti (2007). In fact, by developing an analysis of the effect of informed finance on technological change, they show that banks information promotes an increase in the adoption of new technologies but does not affect R&D investments. In particular, by using similar data from the Capitalia survey, Herrera and Minetti find that the more intense relationship lending is as measured by the length of the relationship between the firm and its main bank the more likely firms feature product or process innovations. However, the length of the credit relationship does not play a significant role for R&D investments. Moreover, the above argument also explains why the main source of financing for R&D investments is internal funds, while, on the contrary, bank lending is considerably less important. In the absence of developed venture capital or private equity markets, as in Italy, R&D expenditures are forced to rely heavily on internal finance and/or public subsidies. In addition, this evidence highlights the disconnect between innovation outputs and formal R&D for many firms.

5.3 Robustness analysis

In this section we apply some robustness checks to accompany our main results.[4] Based on the approach in Guiso et al. (2004a), we investigate whether local banking development affect firm growth; particularly if there are differential effects for our sample of R&D performing firms. Consistent with our previous findings on the irrelevance for R&D expenditures, we might expect that banking development does not play an important role for the growth of R&D firms. In

[4]A major extension to consider is to account for the role of unobserved heterogeneity using panel data analysis. There are some limitations which do not allow for this analysis. Given the short panel component T=3 the econometric model would have to be simple. With addition of future waves would allow for the implementation for a panel component. There is a large amount of attrition with the Capitalia survey. The attrition bias that would be introduced might outweigh the efficiency gains of accounting for unobserved heterogeneity (fixed- effects). Another dimension that could be explored is to consider the role of sample selection or the selection effect of undergoing R&D investment. A Heckman sample selection model has been estimated; the coefficients are similar to the IV and OLS case, therefore, we do not report the results for brevity, but they are available upon request.

table 6 we report the OLS and IV estimates for the specification used in Guiso et al. (2004a), with the same set of instruments. In contrast to our prior, the impact of banking development on firm growth is statistically significant and is more pronounced in the case of small firms with fewer than 50 employees. Overall, the effects of banking development matter more for R&D performing firms.

Our finding on the role of local access to credit for the *growth* of R&D firms extends the analysis on the importance of banking development for firm growth as in Guiso et al. (2004a). The only discriminant found in the literature is the size of the firm, with large firms less exposed to local banking development. We have found another discriminant for the importance of banking development more related to the type of activity of the firm.

Now we perform a robustness analysis on the local R&D intensity of HES. In table 7 we report a number of robustness checks for the estimates reported in table 5, column 4, where the dependent variable is the amount of R&D expenditure. First, we divide the sample of R&D firms into two subsamples: greater than 500 employees and less than or equal to 500 employees. As it is possible to see from columns 1 and 2, the impact of the local R&D intensity of HES is stronger for the subsample of big firms. This finding is also confirmed by the interaction analysis reported in columns 3-5. The coefficient for the interaction with small firms (with less than 50 employees) is not statistically significant. While the coefficients for the interactions with high-tech firms and big firms (with more than 500 employees) are statistically significant.

6 Conclusion

The empirical evidence of this paper confirms the significant role of university R&D in enhancing or crowding-in R&D expenditures in the business sector found in the literature. In particular, novel findings show that the regional R&D intensity of the higher education sector plays a key role in increasing the probability of local cooperation between universities and firms and thereby stimulates firms R&D expenditures locally through indirect effects (or knowledge spillovers). In addition, we estimate a direct effect from firm-university cooperation leading to roughly a 30-percentage point increase in firm R&D investment.

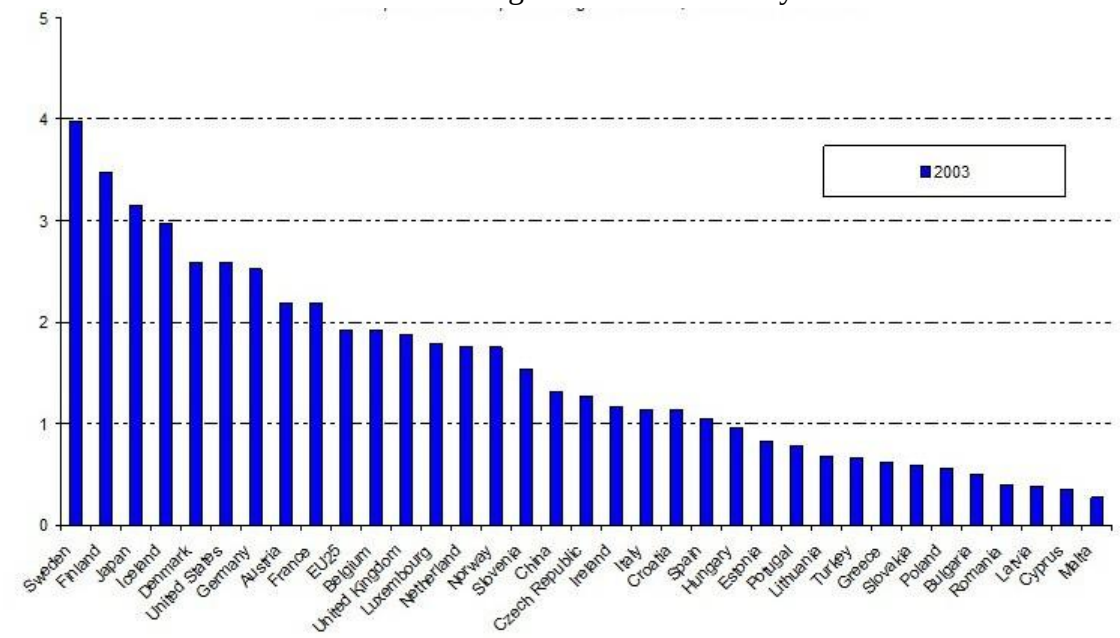
In line with previous evidence, we do not find any effect from local financial development

on firm R&D. Due to the high degree of risk and the complexity of evaluating future prospects of activities related to innovation, banking intermediaries are not ideal for financing R&D expenditures. This fact increases the probability of firms, especially high-tech ones, of being credit constrained independently of the degree of banking development. In contrast to our result on the irrelevance of banking development for R&D expenditures, the analysis of Guiso, Sapienza, and Zingales (2004a) is extended by showing that for R&D firms the effect of local banking development on growth matters and is more pronounced for smaller firms.

These findings have important policy implications. First, our findings suggest also an important role for the government in promoting industrial research not only through demand for subsidies to firms, but also by means of improving the spatial concentration of financial resources for basic research. In particular, strengthening basic research in order to reflect more effectively the needs of the local innovation system could improve local university-industry cooperation as well as lead to a greater transmission of knowledge spillovers thereby stimulating R&D expenditures by the local business sector both through direct and indirect channels.

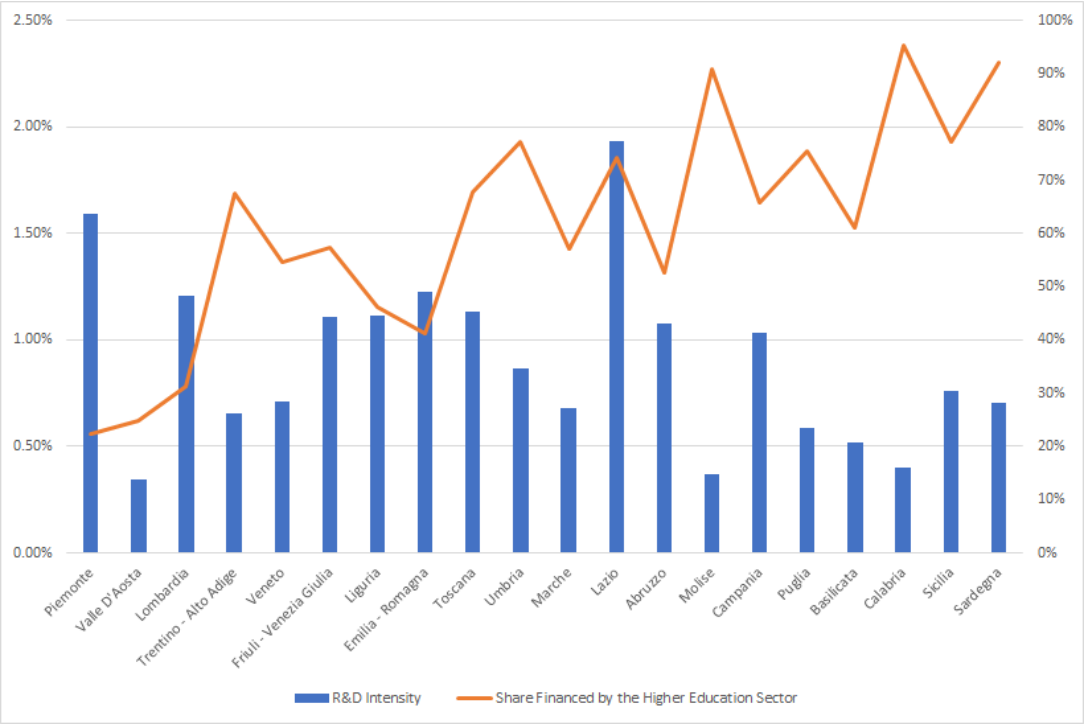
Second, the fact that local banking development does not affect firms R&D expenditures clarifies why even in the more financially developed regions, with relatively easier access to bank credit, firm R&D intensity can be low. In the absence of developed venture capital or private equity markets, such as the case in Italy, it implies R&D expenditures rely heavily on internal finance and on transfers and fiscal subsidies. Thus, there might be a fundamental role for the government in promoting an active specialized capital market, more ideal for financing firms R&D activities.

Figure 1: R&D Intensity



Note: R&D expenditure as percentage of GDP. Source: EUROSTAT.

Figure 2: R&D Intensity in Italian Regions; 2003 statistics



Note: R&D intensity measured by expenditure as percentage of GDP (LHS). Source: ISTAT.

Table 1: Firm R&D Intensity

	1995-1997	1998-2000	2001-2003
Number of firms	4490	4603	4171
Share with R&D > 0			
Overall	29.5	34.3	39.8
Small	20.4	27.4	29.9
Medium	43.1	56.9	48.8
Large	65.1	63.3	68.9
Less than 10 years	29.0	33.0	32.6
10 to 25	26.9	32.6	38.0
Over 25	32.8	36.6	42.7
High Tech	45.2	62.6	59.5
Share of firms conditional on R&D > 0			
University Cooperation	11.7	12.4	13.8
Fiscal incentives	4.8	13.6	10.0
Public transfers	10.0	15.5	18.8
R&D Intensity			
Firms without ties to university, incentives	1.3	2.5	1.2
University cooperation	2.2	3.5	2.0
Fiscal incentives or public transfers	3.3	3.6	3.0

Note: Number of observations refers to firms that have answered the question on R&D in the survey questionnaire. The averages are calculated over each three-year period of the survey. Intensities are R&D ratios relative to production. R&D expenditure is deflated with a weighted average of the hourly earnings in manufacturing index and the aggregate business investment price index, where the weights used are respective 0.9 and 0.1 (see Parisi, Schiantarelli, and Sembenelli (2006)) . Production is computed as the sum of sales, capitalized costs and the change in work-in-progress and in finished goods inventories, with all variables deflated with the appropriate production price index.

Table 2: Sources of Finance for Investments in R&D

	1995-1997	1998-2000	2001-2003
Number of Firms	1489	1645	1828
Sources of finance (% of expenditure)			
New Equity	1.6	1.3	0.7
Internal Funds	82.0	78.8	80.8
Fiscal Incentives	1.7	4.7	2.8
Public Transfers	3.2	5.6	5.8
Bank Debt	9.3	8.3	8.6
Unsubsidized	5.4	5.3	5.4
Subsidized	3.9	3.0	3.2
Importance of bank finance and internal funds			
Share of firms with 100 % internal finance	71.0	63.3	63.0
Share of firms with no bank finance	84.3	87.3	86.0
Share of firms with 100 % unsubsidized bank finance	2.6	2.9	2.3
Share of firms with 100 % subsidized bank finance	1.1	1.2	1.0

Note: Number of observations refers to firms that have answered the question on the sources of finance for R&D spending in the survey questionnaire. Averages are calculated over each three-year period of the survey.

Table 3: The effect of local R&D intensity of higher education sector on university-industry cooperation

	OLS	IV	IV Probit
	(1)	(2)	(3)
Local R&D intensity of HES	0.456*	0.159*	0.716*
	0.25	0.07	0.34
Public transfers	1.077***	0.309***	1.074***
	0.07	0.03	0.07
Fiscal incentives	0.789***	0.206***	0.791***
	0.15	0.05	0.16
Age	0.082*	0.015*	0.083*
	0.04	0.01	0.04
Size	0.241***	0.053***	0.242***
	0.03	0.01	0.03
Per capita GDP	-0.009	-0.001	-0.008
	0.01	0.00	0.01
Social Capital	-1.674	-0.487	-2.124
	1.15	0.30	1.41
South	-0.204	-0.073	-0.328
	0.22	0.06	0.27
Number of Observations	3949	3949	3949
Log likelihood / R^2	-1436.88 / 0.127	0.065	1340.62
Hansen/Sargan (χ^2 p-value)		.609	
Wald (χ^2 p-value)			0.224

Notes: Pooled regressions for three waves covering the period 1995-2003. The left-hand variable is a dummy equal to 1 if the firm cooperates with university on R&D. The regional R&D intensity of the higher education sector is instrumented with a set of variables that describe the regional structure of the higher education sector relevant for knowledge spillovers: the number of universities with faculty of science and the number of total universities in the region for 1995. The Hansen-Sargan test is a test of overidentifying restrictions. The joint null hypothesis is that the instruments are valid instruments, i.e., uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation. Under the null, the test statistic is distributed as χ^2 in the number of overidentifying restrictions. Per capita GDP is the log of per capita net disposable income in the province in 1991. Public transfers, fiscal incentives represent share of R&D expenditure financing. Social capital is measured by average voter turnout at the province level for all referenda in the period between 1946 and 1987 (Guiso, Sapienza, and Zingales (2004b)). South is a dummy for regions of southern Italy, with Lazio excluded. Firm size is the log of the number of employees. Firm's age is the log of the number of years from its birth. Local R&D intensity of higher education sector is the regional R&D expenditure as percentage of GDP in 1995. All regressions include a constant, industry and time dummies. Robust standard errors, reported in brackets, are adjusted for regional clustering. (*),(**), and (***) are coefficient significant at 10, 5, and 1 percent.

Table 4: Treatment effect of university-cooperation on R&D expenditures

		NN		KM	
	OLS	ATE	ATT	ATE	ATT
Effect of university cooperation (in logs)	0.312	0.292	0.326	0.401	0.293
	(0.076)	(0.110)	(0.143)	(0.096)	(0.113)

Note: Numbers displayed are coefficient estimates based on dependent variable of log R&D expenditures with standard errors using 250 bootstrapped replications. Standard errors are in parentheses. Propensity score matching based on NN (nearest-neighbor) and KM (kernel). In percentage effects measured based on ratio of coefficient to average log R&D expenditure for entire sample (ATE) and treatment sample (ATT). N= 3523

Table 5: Determinants of firm R&D: influence of local factors

	OLS	IV	OLS	IV	IV
	(1)	(2)	(3)	(4)	(5)
Local banking development	-0.112	0.574			0.760
	0.36	0.62			0.51
Local R&D intensity of HES			0.470***	0.974*	0.846***
			0.10	0.39	0.23
Public transfers	0.901***	0.901***	0.902***	0.904***	0.904***
	0.12	0.11	0.12	0.11	0.11
Fiscal incentives	0.733***	0.731***	0.741***	0.750***	0.746***
	0.12	0.11	0.12	0.12	0.12
Internally financed growth	0.582*	0.568*	0.605*	0.632*	0.611*
	0.28	0.28	0.28	0.27	0.26
Size	0.907***	0.906***	0.910***	0.913***	0.911***
	0.01	0.01	0.01	0.01	0.01
Per capita GDP	0.014	0.016	0.015	0.016*	0.017*
	0.01	0.01	0.01	0.01	0.01
Social capital	0.458	0.111	-0.397	-1.250	-1.415
	1.13	1.05	0.97	1.18	1.15
South	-0.161	0.010	-0.358*	-0.598*	-0.347
	0.21	0.31	0.15	0.27	0.21
Number of Observations	3119	3119	3119	3119	3119
R^2	0.447	0.449	0.401	0.401	0.403
Hansen/Sargan χ^2 p-value			0.401	0.479	0.204

Notes: Pooled regressions for three waves covering the period 1995-2003. The left-hand variable is the log of R&D expenditure deflated with a weighted average of the hourly earnings in manufacturing index and the aggregate business investment price index, where the weights used are respectively 0.9 and 0.1 (see Parisi, Schiantarelli, and Sembenelli (2006)). There are two set of instruments that are utilized for IV estimation. First, in the IV estimation in column (3), the indicator of financial development is instrumented with a set of variables that describes the banking market as of 1936 (see Guiso, Sapienza, and Zingales (2004a)). Second, in the IV estimation in column (4), the regional R&D intensity of the higher education sector is instrumented with a set of variables that describe the regional structure of the higher education sector relevant for knowledge spillovers: the per capita number of universities with faculty of science and the per capita number of total universities in the region for 1995. In the IV estimation in column (5) both set of instruments are utilized. The Hansen-Sargan test is a test of overidentifying restrictions. The joint null hypothesis is that the instruments are valid instruments, i.e., uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation. Under the null, the test statistic is distributed as χ^2 in the number of overidentifying restrictions. Banking development is the indicator of local financial development computed by Guiso, Sapienza, and Zingales (2004a). Per capita GDP is the per capita net disposable income in the province in 1991. Social capital is measured by average voter turnout at the province level for all referenda in the period between 1946 and 1987 (see Guiso, Sapienza, and Zingales (2004b)). South is a dummy for regions of southern Italy, with Lazio excluded. The maximum rate of growth internally financed is max g = ROA/(1-ROA), where ROA is the return on assets. Firm size is the log of the number of employees. Local R&D intensity of higher education sector is the regional R&D expenditure as percentage of GDP in 1995. All regressions include a constant, industry and time dummies. Robust standard errors, reported in brackets, are adjusted for regional clustering. (*),(**), and (***) are coefficient significant at 10, 5, and 1 percent.

Table 6: The effects of banking development on the growth of R&D firms

	Full sample			R&D firms only		
			Small firms			Small firms
	OLS	IV	IV	OLS	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Local banking development	0.016	0.049**	0.059**	0.023	0.068**	0.131***
	0.02	0.02	0.02	0.02	0.03	0.05
Interacted with R&D	0.024***	0.023***	0.022***			
	0.01	0.01	0.01			
Internally financed growth	0.036	0.035	0.044	0.059	0.059*	0.083*
	0.02	0.02	0.03	0.03	0.03	0.05
Size	0.004***	0.004***	0.016***	0.002	0.002	0.007
	0.00	0.00	0.00	0.00	0.00	0.01
Per capita GDP	-0.001***	-0.001***	-0.001**	-0.001***	-0.001***	-0.002*
	0.00	0.00	0.00	0.00	0.00	0.00
Social Capital	0.039	0.019	0.032	-0.025	-0.049	-0.156
	0.04	0.05	0.06	0.06	0.07	0.10
Judicial efficiency	-0.003	-0.003	-0.004	-0.003	-0.003	-0.004
	0.00	0.00	0.00	0.00	0.00	0.00
South	0.024**	0.032***	0.031*	0.012	0.024**	0.030
	0.01	0.01	0.02	0.01	0.01	0.03
Observations	8782	8779	5347	3115	3114	1403
Hansen/Sargan χ^2 p-value		0.2638	0.0764		0.4540	0.9661

Notes: Pooled regressions for four waves covering the period 1992-2003. The left-hand variable is the log change in firm production. IV uses as instrument a set of variables that describes the banking market as of 1936 (see Guiso, Sapienza, and Zingales (2004a)). The Hansen-Sargan test is a test of overidentifying restrictions. The joint null hypothesis is that the instruments are valid instruments, i.e., uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation. Under the null, the test statistic is distributed as χ^2 in the number of overidentifying restrictions. The (1) column includes first two waves, (2) column is last two waves, (3) column is first three waves, and (4) column is all four waves. Banking development is the indicator of local financial development computed by Guiso, Sapienza, and Zingales (2004a). Per capita GDP is the per capita net disposable income in the province in 1991. Social capital is measured by average voter turnout at the province level for all referenda in the period between 1946 and 1987 (see Guiso, Sapienza, and Zingales (2004b)). South is a dummy for regions of southern Italy, with Lazio excluded. The maximum rate of growth internally financed is max g = ROA/(1-ROA), where ROA is the return on assets. Firm size is the log of the number of employees. Local R&D intensity of higher education sector is the regional R&D expenditure as percentage of GDP in 1995. All regressions include a constant, industry and time dummies. Robust standard errors, reported in brackets, are adjusted for regional clustering. (*),(**), and (***) are coefficient significant at 10, 5, and 1 percent.

Table 7: Robustness for Local R&D intensity of higher education sector

	Subsamples		Interaction with		
	> 500 EMP	≤ 500 EMP	Hi-tech	Big Firms	Small Firms
	(1)	(2)	(3)	(4)	(5)
Local R&D intensity of HES	2.056*	0.797*	0.939*	0.887*	0.871*
	0.87	0.41	0.41	0.39	0.43
Hi-tech firms			0.385*		
			0.22		
Large firms				0.633**	
				0.22	
Small firms					0.151
					0.18
Public transfers	0.917	0.924***	0.907***	0.911***	0.906***
	0.85	0.10	0.11	0.11	0.11
Fiscal incentives	0.384	0.797***	0.734***	0.754***	0.755***
	0.90	0.12	0.12	0.12	0.12
Internally financed growth	4.115***	0.318	0.663*	0.636*	0.623*
	1.07	0.27	0.27	0.27	0.27
Size	1.246***	0.857***	0.912***	0.883***	0.932***
	0.07	0.02	0.01	0.02	0.02
Per capita GDP	0.015	0.014**	0.014*	0.014*	0.015*
	0.03	0.01	0.01	0.01	0.01
Social capital	-5.216*	-0.652	-1.367	-1.170	-1.185
	2.48	1.22	1.24	1.17	1.17
South	-1.331**	-0.506	-0.645*	-0.585*	-0.581*
	0.49	0.26	0.29	0.27	0.27
Number of Observations	212	2904	3119	3119	3119
R^2	0.314	0.293	0.402	0.403	0.404
Hansen/Sargan χ^2 p-value	0.424	0.448	0.455	0.449	0.428

Notes: Pooled regressions for three waves covering the period 1995-2003. The left-hand variable is the log of R&D expenditure deflated with a weighted average of the hourly earnings in manufacturing index and the aggregate business investment price index, where the weights used are respectively 0.9 and 0.1 (see Parisi, Schiantarelli, and Sembenelli (2006)). The regional R&D intensity of the higher education sector is instrumented with a set of variables that describe the regional structure of the higher education sector relevant for knowledge spillovers: the per capita number of universities with faculty of science and the per capita number of total universities in the region for 1995. The Hansen-Sargan test is a test of overidentifying restrictions. The joint null hypothesis is that the instruments are valid instruments, i.e., uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation. Under the null, the test statistic is distributed as χ^2 in the number of overidentifying restrictions. The (1) column includes firms with at least 500 employees, (2) column is less than or equal to 500 employees, (3) column includes interaction with an indicator of high technology firms, (4) column is interaction with an indicator of firm with at least 500 employees, and (5) is a combination of all interactions. Local R&D intensity of higher education sector is the regional R&D expenditure as percentage of GDP in 1995. Per capita GDP is the per capita net disposable income in the province in 1991. Social capital is measured by average voter turnout at the province level for all referenda in the period between 1946 and 1987 (see Guiso, Sapienza, and Zingales (2004b)). South is a dummy for regions of southern Italy, with Lazio excluded. The maximum rate of growth internally financed is max g = ROA/(1-ROA), where ROA is the return on assets. Firm size is the log of the number of employees. All regressions include a constant, industry and time dummies. Robust standard errors, reported in brackets, are adjusted for regional clustering. (*),(**), and (***) are coefficient significant at 10, 5, and 1 percent.

Table A.1: Descriptive Statistics

	MAX	MIN	MEDIAN	MEAN	1ST PERC	99TH PERC	ST DEV
South	1	0	0	0.12	0	1	0.32
Local banking development	0.59	0	0.44	0.43	0.03	0.59	0.12
LOG firm's age	5.26	0	3.14	3.09	1.1	4.67	0.72
LOG (number of employees)	9.39	1.99	3.94	4.18	2.48	7.49	1.21
GDP per capita in 1991 (euro)	23.53	8.32	16.7	16.71	8.65	23.53	3.38
Judicial efficiency	7.47	1.44	2.82	3.14	1.88	6.79	0.89
Social capital	0.92	0.62	0.86	0.85	0.66	0.92	0.05
Maximum growth internally financed	0.59	0	0.08	0.1	0.01	0.4	0.07
Local R&D intensity of higher education sector (perc. of GDP)	1.29	0.01	0.28	0.38	0.23	1.29	0.18
Bank branches per 10000 inhabitants in the region in 1936	5.31	0.57	2.22	2.55	0.83	5.31	1.18
Share of bank branches owned by local banks in 1936	0.97	0.46	0.89	0.83	0.51	0.97	0.13
Savings banks per 10000 inhabitants in the region in 1936	0.12	0	0.03	0.03	0	0.1	0.03
Cooperative banks per 10000 inhabitants in the region in 1936	0.22	0	0.06	0.09	0	0.22	0.06
Share of new equity used for financing R&D	1	0	0	0.01	0	0.6	0.1
Share of internal funds used for financing R&D	1	0	1	0.8	0	1	0.34
Share of fiscal incentives used for financing R&D	1	0	0	0.03	0	0.86	0.13
Share of public transfers used for financing R&D	1	0	0	0.05	0	1	0.17
Share of bank debt used for financing R&D	1	0	0	0.06	0	1	0.2
Share of bank debt subsidized used financing R&D	1	0	0	0.03	0	1	0.15
Cooperation with university on R&D	1	0	0	0.15	0	1	0.36
# of universities with faculty of science in the region in 1995	6	1	4	3.54	1	6	1.06
Total number of universities in the region in 1935	9	1	4	4.95	1	9	2.79
Rate of growth of production	0.97	-0.82	0.02	0.03	-0.34	0.45	0.14
Log R&D expenditures	14.86	-2.34	6.69	6.8	2.96	11.76	1.77
R&D intensity (perc. of production)	21.87	0	0.99	1.99	0.02	14.56	2.8

Note: Descriptive statistics for the data of the Capitalia Survey on Italian SME, for three waves covering the period 1995-2003.

Conclusion Générale

In this thesis, we study the effects of innovation on firm performance in Chapters I-III as well as the determinants of firm-level investment in R&D in Chapter IV. In all cases, we derive empirical findings using micro data derived from firm surveys and a combination of econometric methods. In Chapters I and II, we study the effects of innovation on firm productivity and subsequently investigate the trade-offs between innovation and employment using a large sample of firms from a cross-country perspective in the developing world. In chapter 3, we study the role of innovation and growth for a representative sample of startups in the United States where we make use of the longitudinal nature of the data to observe firm growth before and after innovation. In Chapter 4, we study the role of university-firm cooperation on R&D projects as having a crowding-in effect on firm R&D investment for a representative cross-sectional sample of manufacturing firms from Italy. Our contribution falls in the field of industrial economics with a connection to broader macroeconomic growth.

We highlight some of the key findings and their relevance for policy as follows. When benchmarking innovation rates across a set of developing countries, we found that the rate of innovation when defined as imitation declines as income per capita increases; but the reverse is the case when innovation is defined as radical. Incidence of R&D increases as income per capita increases and similarly is most often associated with radical innovation. While the returns to innovation are generally positive, we find substantially larger effects for product innovations that are deemed radical as compared to imitation. Further, the returns to process innovation are generally lower than product innovation; though they are found to be just as large when excluded to automation. In contrast to our hypothesis that returns to innovation might be greatest in countries farthest from the technological frontier, the lack of higher returns for either imitation or radical innovation in these countries suggest that policy could have a greater focus on fostering complementary firm capabilities.

While innovation is seen as an important driver to boost firm-level productivity and ultimately increase income per capita, there is a short-term tension with the possible displacement of labor. Untangling the short-run effect on employment from firm innovation activities, we generally find that product innovation, when successful in bringing additional sales to the firm, has a positive direct impact on employment in the short-run. However, the extent to which new sales from product innovation cause additional employment, however, is directly related to the impact on efficiency resulting from the innovation process, which seems to vary with the income level of the country. Interestingly, the efficiency parameter is estimated to be lower in lower income countries resulting

in the tendency for innovation to be labor enhancing. However, our results are agnostic about the potential effects from innovation on the skill composition of the firm and the potential displacement on unskilled workers. As a result, innovation policy can be an effective lever to promote employment in the short-run, particularly when combined with job skills training.

Using a longitudinal representative sample of start-up firms in the United States over the period 2004-2011, we investigated the role of innovation and finance in affecting firm performance in terms of both survival and growth. The literature has highlighted the role of successful, high-growth newly launched firms in contributing to aggregate productivity and employment gains. However, many new firms face poor survival prospects augmented by financial frictions and difficulty in raising external finance. We find that firms in the high-tech sector display an inverted U-shaped hazard rate and initial financing levels positively affect survival; but the later effect is reversed during a time of recession. The introduction of product innovations midway through the firm's observed lifespan is associated with positive sales growth from a before after perspective but we do not find any effect on employment growth. Our results are informative for policymakers in terms of highlighting not only the importance of financing but also the role of firm quality proxied by innovation outputs.

In our study on manufacturing firms in Italy, we confirm the significant role of university R&D in enhancing or crowding-in R&D expenditures in the business sector found in the literature. In particular, regional R&D intensity of the higher education sector both increases the probability of local cooperation between universities and firms and stimulates firm R&D expenditures locally potentially through indirect effects (or knowledge spillovers). In addition, we estimate a direct effect from firm-university cooperation leading to roughly a 30-percentage point increase in firm R&D investment which is roughly equal in magnitude to the average R&D tax credit. Our findings suggest an important role for policy in promoting industrial research not only through demand for subsidies to firms, but also by dedicating financial resources for basic research undertaken by the university sector that aligns with the needs of the local innovation system.

While the quantitative methods we employ are simple, the novelty of our approach lies in making use of newly introduced measures of innovation based on the Oslo Manual guidelines in two different contexts yet to be studied previously: established firms in developing countries, and newly launched startups in an advanced economy. In both cases, our work illustrates the challenge of using such innovation questions for research. Survey data collected from the field is often plagued with measurement challenges. In the case of newly established firms, care is also needed when dealing

with the nature of the variables and corresponding survey responses for new firms that are just getting off the ground. In the final chapter, our contribution stems from providing empirical evidence on an under-studied policy lever for incentivizing and promoting R&D investment among the private sector by encouraging cooperation with the university sector. Our findings contribute to the industrial economics literature towards a better understanding of the drivers and heterogeneity of firm performance as well as highlighting the role of and importance of innovation.

Some major limitations of our work lie in the limited time frame with which we observe the firm that in some cases exclude more advanced panel-data methods, the focus on average effects that could mask heterogenous effects as well as outstanding limitations with the nature of our innovation variables. At the same time, the multi-year process of innovation calls for observing the firm over a sufficiently long-time frame to adequately assess short-run vs. long-run effects. While the enhanced nature of the questions on innovation, such as the degree of novelty, offer avenues to exploit heterogeneity and establish patterns in the data, the subjective nature of the data pose risks of misinterpretation and misclassification. More recently, an ongoing effort in survey data collection is exploring the role of technology adoption. Future research could explore these methodological aspects in greater detail and make use of alternate datasets.

Bibliographie

ACEMOGLU, D., U. AKCIGIT, N. BLOOM, and W. KERR (2013): "Innovation, Reallocation And Growth," Working Papers 13-23, Center for Economic Studies, U.S. Census Bureau.

ADAMS, J. D. (2002): "Comparative localization of academic and industrial spillovers," *Journal of Economic Geography*, 2(3), 253–278.

ADAMS, J. D., E. P. CHIANG, and K. STARKEY (2001): "Industry-University Cooperative Research Centers," *The Journal of Technology Transfer*, 26(1-2), 73–86.

AGHION, P., and P. HOWITT (1992): "A Model of Growth through Creative Destruction," *Econometrica*, 60(2), 323–51.

AGRAWAL, A., and I. M. COCKBURN (2002): "University Research, Industrial R&D, and the anchor Tenant Hypothesis," NBER Working Papers 9212, National Bureau of Economic Research, Inc.

AKCIGIT, U., and W. R. KERR (2010): "Growth Through Heterogeneous Innovations," NBER Working Papers 16443, National Bureau of Economic Research, Inc.

ARORA, A., S. BELENZON, A. PATACCONI, AND J. SUH (2019): "The Changing Structure of American Innovation: Some Cautionary Remarks for Economic Growth," in Innovation Policy and the Economy, Volume 20, NBER Chapters, pp. 39–93. National Bureau of Economic Research, Inc.

ARRIGHETTI, A., and M. VIVARELLI (1999): "The Role of Innovation in the Post entry Performance of New Small Firms: Evidence from Italy," *Southern Economic Journal*, 65(4), 927–939.

AUTOR, D. H., and D. DORN (2013): "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103(5), 1553–97.

BAUM, C. F., H. LOOF, P. NABAVI, and A. STEPHAN (2015): "A New Approach to Estimation of the R&D-Innovation-Productivity Relationship," Boston College Working Papers in Economics 876, Boston College Department of Economics.

BELL, M., and K. PAVITT (1997): "Technological accumulation and industrial growth: contrasts be- tween developed and developing countries," *Technology, globalisation and economic performance*, 83137, 83–137.

BENAVENTE, J. M., and R. LAUTERBACH (2008): "Technological innovation and employment: complements or substitutes?" *The European Journal of Development Research*, 20(2), 318–329.

BENFRATELLO, L., F. SCHIANTARELLI, and A. SEMBENELLI (2008): "Banks and innovation: Microeconometric evidence on Italian firms," *Journal of Financial Economics*, 90(2), 197–217.

BERGER, A. N., and G. F. UDELL (1992): "Some Evidence on the Empirical Significance of Credit Rationing," *Journal of Political Economy*, 100(5), 1047–77.

BERGER, A. N., G. CERQUEIRO, and M. F. PENAS (2014): "Market Size Structure and Small Business Lending: Are Crisis Times Different from Normal Times?," *Journal of Finance*.

BIANCHI, P., and L. RAMACCIOTTI (2005): *Research and Technological Innovation: The Challenge for a New Europe* chap. Relationships between Universities, research centers and district firms: the Italian case, pp. 265–283. Physica-Verlag, Heidelberg.

BLOOM, N., M. SCHANKERMAN, and J. V. REENEN (2005): "Identifying technology spillovers and product market rivalry," *Federal Reserve Bank of San Francisco Proceedings*.

BLOOM, N., VAN REENEN, J., and WILLIAMS, H. 2019. "A Toolkit of Policies to Promote Innovation." *Journal of Economic Perspectives*, 33 (3): 163-84.

BONNET, J., and R. CRESSY (2016): "Title: An encompassing model and test of the Evans-Jovanovic credit constraints hypothesis," Economics Working Paper Archive (University of Rennes 1 & University of Caen) 2016-01, Center for Research in Economics and Management (CREM), University of Rennes 1, University of Caen and CNRS.

BROWN, J. D., and J. S. EARLE (2013): "Do SBA Loans Create Jobs?," IZA Discussion Papers 7544, Institute for the Study of Labor (IZA).

BROWN, J. D., and J. S. EARLE (2010): "Entry, Growth, and the Business Environment: A Comparative Analysis of Enterprise Data from the U.S. and Transition Economies," Working Papers 10-20, Center for Economic Studies, U.S. Census Bureau.

BROWN, J. D., J. S. EARLE, M. J. KIM, AND K. M. LEE (2019): "Immigrant Entrepreneurs and Innovation in the US High-Tech Sector," in The Roles of Immigrants and Foreign Students in US Science, Innovation, and Entrepreneurship, NBER Chapters, pp. 149–171. National Bureau of Economic Research, Inc.

BROWN, J., J. EARLE, and D. LUP (2005): "What Makes Small Firms Grow? Finance, Human Capital, Technical Assistance, and the Business Environment in Romania," *Economic Development and Cultural Change*, 54(1), 33–70.

BRYNJOLFSSON, E., and A. MCAFEE (2012): *Race Against the Machine: How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy*. Digital Frontier Press.

CALVINO, F. and M.E. VIRGILLITO (2018): "The innovation-employment nexus: a critical survey of theory and empirics", *Journal of Economic Surveys*, 32(1), 83-117.

CAMERON, A., and P. K. TRIVEDI (2005): *Microeconometrics: Methods and Applications*. Cambridge University Press.

CASSIMAN, B., and R. VEUGELERS (2006): "In Search of Complementarity in Innovation Strategy: Internal R& D and External Knowledge Acquisition," *Management Science*, 52(1), 68–82.

CASTILLO, V., A. MAFFIOLI, S. ROJO, and R. STUCCHI (2014): "The effect of innovation policy on SMEs employment and wages in Argentina," *Small Business Economics*, 42(2), 387–406.

CI, W., J. GALDO, M. VOIA, and C. WORSWICK (2015): "Wage returns to mid-career invest- ments in job training through employer supported course enrollment: evidence for Canada," *IZA Journal of Labor Policy*, 4(1), 1–25.

CIRERA, X. (2015): "Catching Up to the Technological Frontier?," World Bank Other Operational Studies 21684, The World Bank.

CIRERA, X. and L. SABETTI (2016) "The Effects of Innovation on Employment in Developing Countries: Evidence from Enterprise Surveys" *Policy Research Working Paper* 7775. World Bank, Washington, DC.

CIRERA, X. & L. SABETTI, (2019). "The effects of innovation on employment in developing countries: evidence from enterprise surveys," *Industrial and Corporate Change*, Oxford University Press, vol. 28(1), pages 161-176.

CIRERA, X., and S. MUZI (2016): "Measuring firm-level innovation using short questionnaires: evidence from an experiment," Policy Research Working Paper Series 7696, The World Bank.

CIRERA, X., and W. F. MALONEY (2017): *The Innovation Paradox*, no. 28341 in World Bank Publications. The World Bank.

CIRERA, X., L. IACOVONE, and L. SABETTI (2015): "Catching Up to the Technological Frontier? Understanding Firm-Level Innovation in Developing Countries," Discussion paper, The World Bank.

COHEN, W M., (2010), "Fifty Years of Empirical Studies of Innovative Activity and Performance", Chapter 4, p. 129-213, *Handbook of the Economics of Innovation*, Elsevier.

COHEN, W. M., and D. A. LEVINTHAL (1989): "Innovation and Learning: The Two Faces of R&D," *Economic Journal*, 99(397), 569–96.

COHEN, W. M., R. R. NELSON, and J. P. WALSH (2000): "Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)," NBER Working Papers 7552, National Bureau of Economic Research, Inc.

COSH, A., D. CUMMING, and A. HUGHES (2009): "Outside Entrepreneurial Capital," *Economic Journal*, 119(540), 1494–1533.

CREPON, B., E. DUGUET, and J. MAIRESSE (1998): "Research, Innovation And Productivity: An Econometric Analysis At The Firm Level," *Economics of Innovation and New Technology*, 7(2), 115–158.

CRESPI, G. and E. TACSIR (2013) "Effects of innovation on employment in Latin America," *MERIT Working Papers* 001, United Nations University - Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT).

CRESPI, G., E. ARIAS-ORTIZ, E. TACSIR, F. VARGAS, and P. ZUIGA (2014): "Innovation for economic performance: the case of Latin American firms," *Eurasian Business Review*, 4(1), 31–50.

CRISCUOLO, C. (2009): "Innovation and Productivity," *Innovation in Firms*, pp. 111–138.

DACHS, B., and B. PETERS (2014): "Innovation, employment growth, and foreign ownership of firms," *Research Policy*, 43(1), 214–232.

DECKER, R. A., J. HALTIWANGER, R. S. JARMIN, and J. MIRANDA (2016): "Declining Business Dynamism: What We Know and the Way Forward," *American Economic Review*, 106(5), 203–07.

DESROCHES, D., T. BARTON, J. BALLOU, F. POTTER, Z. ZHAO, B. SANTOS, and J. SEBASTIAN (2007): "Kauffman Firm Survey (KFS) Baseline Methodology Report," Mathematica Policy Research Reports 5568, Mathematica Policy Research.

DOSI, G., S. LECHEVALIER, AND A. SECCHI (2010): "Introduction: Interfirm heterogeneity—nature, sources and consequences for industrial dynamics," *Industrial and Corporate Change*, 19(6), 1867–1890.

DRUKKER, D. M. (2011): "Estimating the parameters of simultaneous-equations models with the SEM command in Stata 12," Italian Stata Users' Group Meetings 2011 01, Stata Users Group.

FAGERBERG, J., D. C. MOWERY, and R. R. NELSON (2006): *The Oxford handbook of innovation*. Oxford university press.

FELLER, I. (1990): "Universities as engines of R&D-based economic growth: They think they can," *Research Policy*, 19(4), 335–348.

FOSTER, L., J. HALTIWANGER, and C. KRIZAN (2002): "The Link Between Aggregate and Micro Productivity Growth: Evidence from Retail Trade," NBER Working Papers 9120, National Bureau of Economic Research, Inc.

FREY, C. B., and M. A. OBORNE (2016): "The future of employment: how susceptible are jobs to computerisation," *Technological Forecasting and Social Change*, Volume 114, January 2017, Pages 254-280.

GOEDHUYS, M., N. JANZ, and P. MOHNEN (2008): "What drives productivity in Tanzanian manufacturing firms: technology or business environment?," *The European Journal of Development Research*, 20(2), 199–218.

GONI, E., and W. F. MALONEY (2014): "Why don't poor countries do R&D ?," Policy Research Working Paper Series 6811, The World Bank.

GRIFFITH, R., E. HUERGO, J. MAIRESSE, and B. PETERS (2006): "Innovation and Productivity Across Four European Countries," *Oxford Review of Economic Policy*, 22(4), 483–498.

GUISO, L. (1998): "High-tech firms and credit rationing," *Journal of Economic Behavior & Organization*, 35(1), 39–59.

GUISO, L., P. SAPIENZA, and L. ZINGALES (2004a): "Does Local Financial Development Matter?," *The Quarterly Journal of Economics*, 119(3), 929–969.

GUISO, L., P. SAPIENZA, and L. ZINGALES (2004b): "The Role of Social Capital in Financial Development," *American Economic Review*, 94(3), 526–556.

HADLOCK, P., D. HECKLER, and J. GANNON (1991): "High-Technology Employment Another View," *Monthly Labor Review*, pp. 26–30.

HALL, B. H. (2011): "Innovation and productivity," Discussion paper.

HALL, B. H., A. N. LINK, and J. T. SCOTT (2003): "Universities as Research Partners," *The Review of Economics and Statistics*, 85(2), 485–491.

HALL, B. H., F. LOTTI, and J. MAIRESSE (2008): "Employment, innovation, and productivity: evidence from Italian microdata," *Industrial and Corporate Change*, 17(4), 813–839.

HALL, B., F. LOTTI, and J. MAIRESSE (2009): "Innovation and productivity in SMEs: empirical evidence for Italy," *Small Business Economics*, 33(1), 13–33.

HALTIWANGER, J., R. S. JARMIN, AND J. MIRANDA (2013): "Who Creates Jobs? Small versus Large versus Young," *The Review of Economics and Statistics*, 95(2), 347–361.

HARRISON, R., J. JAUMANDREU, J. MAIRESSE, AND B. PETERS (2014): "Does innovation stimulate employment? A firm-level analysis using comparable micro-data from four European countries," *International Journal of Industrial Organization*, 35(C), 29–43.

HERRERA, A. M., and R. MINETTI (2007): "Informed finance and technological change: Evidence from credit relationships," *Journal of Financial Economics*, 83(1), 223–269.

HO, D., K. IMAI, G. KING, and E. STUART (2007): "Matching as Nonparametric Pre-processing for Reducing Model Dependence in Parametric Causal Inference," *Political*

Analysis, 15, 199–236.

HOPENHAYN, H. A. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60(5), 1127–50.

HSIAO, C. (2002): *Analysis of Panel Data.* Cambridge University Press, second edition edn.

HSIEH, C.-T., and P. J. KLENOW (2009): "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 124(4), 1403–1448.

HURST, E., and B. W. PUGSLEY (2011): "What do Small Businesses Do?," *Brookings Papers on Economic Activity*, 43(2 (Fall)), 73–142.

HUYNH, K. P., R. J. PETRUNIA, and M. VOIA, 2010. "The Impact Of Initial Financial State On Firm Duration Across Entry Cohorts," *Journal of Industrial Economics*, Wiley Blackwell, vol. 58(3), pages 661-689, September.

HUYNH, K. P., R. J. PETRUNIA, and M. VOIA, 2012a."Initial Financial Conditions, Unobserved Heterogeneity and the Survival of Nascent Canadian Manufacturing Firms," *Managerial and Decision Economics*, John Wiley & Sons, Ltd., vol. 33, pages 109-125, March.

HUYNH, K. P., R. J. PETRUNIA, and M. VOIA (2012b): "Duration of new firms: The role of startup financial conditions, industry and aggregate factors," *Structural Change and Economic Dynamics, Elsevier*, 23(4), 354–362.

HUYNH, K. P., and M. VOIA, 2017. "Mixed proportional hazard models with continuous finite mixture unobserved heterogeneity: an application to Canadian firm survival," *Applied Stochastic Models in Business and Industry*, John Wiley & Sons, vol. 33(1), pages 81-94, January.

HUYNH, K. P., and R. J. PETRUNIA,, 2010. "Age effects, leverage and firm growth," Journal of Economic Dynamics and Control, Elsevier, vol. 34(5), pages 1003-1013, May.

HUYNH, K. P. and ROTONDI, ZENO, R&D Spending and Knowledge Spillovers (July 15, 2009). Available at SSRN: https://ssrn.com/abstract=968666 or http://dx.doi.org/10.2139/ssrn.968666

JACOBS, B., and F. VAN DER PLOEG (2006): "Guide to reform of higher education: a European perspective," *Economic Policy*, 21(47), 535–592.

JAFFE, A. B., M. TRAJTENBERG, and R. HENDERSON (1993): "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *The Quarterly Journal of Economics*, 108(3), 577–98.

KLETTE, T. J., and S. KORTUM (2004): "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 112(5), 986–1018.

KLINE, S. J., and N. ROSENBERG (1986): "An overview of innovation," *The positive sum strategy: Harnessing technology for economic growth*, 14, 640.

KROSZNER, R. S., L. LAEVEN, and D. KLINGEBIEL (2007): "Banking crises, financial dependence, and growth," *Journal of Financial Economics*, 84(1), 187–228.

LENTZ, R., and D. T. MORTENSEN (2008): "An Empirical Model of Growth Through Product Innovation," *Econometrica*, 76(6), 1317–1373.

MAIRESSE, J., and P. MOHNEN (2010): "Using Innovation Surveys for Econometric Analysis," in *Handbook of the Economics of Innovation*, ed. by B. H. Hall, and N. Rosenberg, vol. 2 of *Handbook of the Economics of Innovation*, Chapter 0, pp. 1129–1155. Elsevier.

MINETTI, R., and S. C. ZHU (2011): "Credit constraints and firm export: Microeconomic evidence from Italy," *Journal of International Economics*, 83(2), 109–125.

MIRAVETE, E. J., and J. C. PERNAS (2006): "INNOVATION COMPLEMENTARITY and SCALE OF PRODUCTION -super-," *Journal of Industrial Economics*, 54(1), 1–29.

MOHNEN, P., and B. H. HALL (2013): "Innovation and productivity: An update," Discussion paper.

MUSOLESI, A., and J.-P. HUIBAN (2010): "Innovation and productivity in knowledge intensive business services," *Journal of Productivity Analysis*, 34(1), 63–81.

PARISI, M. L., F. SCHIANTARELLI, and A. SEMBENELLI (2006): "Productivity, innovation and R&D: Micro evidence for Italy," *European Economic Review*, 50(8), 2037–2061.

PARKER, S. C. (2009): *The Economics of Entrepreneurship*, no. 9780521728355 in Cambridge Books. Cambridge University Press.

PETERSEN, M. A., and R. G. RAJAN (2002): "Does Distance Still Matter? The Information Revolution in Small Business Lending," *Journal of Finance*, 57(6), 2533–2570.

PHILLIPS, B. D., and B. A. KIRCHHOFF (1989): "Formation, growth and survival; Small firm dynamics in the U.S. Economy," *Small Business Economics*, 1(1), 65–74.

PHILLIPS, G. M., and A. ZHDANOV (2012): "R&D and the Incentive from Merger and Acquisition Activity," *Review of Financial Studies*, 0(0), 1–45.

POLDER, M., G. VAN LEEUWEN, P. MOHNEN, and W. RAYMOND (2010): "Product, process and organizational innovation: drivers, complementarity, and productivity effects," MPRA Paper 23719, University Library of Munich, Germany.

RABE-HESKETH, S., AND A. SKRONDAL (2012): Multilevel and Longitudinal Modeling Using Stata, 3rd Edition. StataCorp LP, 3rd edn.

ROBB, A. M., and D. T. ROBINSON (2014): "The Capital Structure Decisions of New Firms," *Review of Financial Studies*, 27(1), 153–179.

ROMER, P. M. (1986): "Increasing Returns and Long-run Growth," *Journal of Political Economy*, 94(5), 1002–37.

ROODMAN, D. (2011): "Fitting fully observed recursive mixed-process models with cmp," *Stata Journal*, 11(2), 159–206.

ROSENBAUM, P. (2002): *Observational Studies*, Springer, New York, NY.

ROSENBERG, N. (1990): "Why do firms do basic research (with their own money)?," *Research Policy*, 19(2), 165 – 174.

SCHFER, D., A. STEPHAN, and J. S. MOSQUERA (2015): "Innovation Capabilities and Financing Constraints of Family Firms," Working Paper Series in Economics and Institutions of Innovation 425, Royal Institute of Technology, CESIS - Centre of Excellence for Science and Innovation Studies.

SCHUMPETER, J. (1942) *Capitalism, Socialism, and Democracy*. New York: Harper & Bros.

SIEGEL, D. S., and P. H. PHAN (2004): "Analyzing the Effectiveness of University Technology Transfer: Implications for Entrepreneurship Education," Rensselaer Working Papers in Economics 0426, Rensselaer Polytechnic Institute, Department of Economics.

SOLOW, R. M. (1957): "Technical change and the aggregate production function," *The review of Economics and Statistics*, pp. 312–320.

SPALIARA, M.-E., and S. TSOUKAS (2013): "What matters for corporate failures in Asia? Exploring the role of firm-specific characteristics during the Asian crisis," *Structural Change and Economic Dynamics*, 26(C), 83–96.

STIGLITZ, J. E., and A. WEISS (1981): "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 71(3), 393–410.

STOCK, J., AND M. YOGO (2005): Testing for Weak Instruments in Linear IV Regression pp. 80–108. Cambridge University Press, New York.

SUTTON, J. (2005): "Competing in capabilities: an informal overview," *First Development Economic Lectures, Development Economics Vice Presidency (DEC), World Bank*.

SYVERSON, C. (2011): "What Determines Productivity?," *Journal of Economic Literature*, 49(2), 326–65.

TABARROK, A. (2002): "Patent Theory versus Patent Law," *The B.E. Journal of Economic Analysis & Policy*, 0(1), 9.

TETHER, B. (2003): "The sources and aims of innovation in services: Variety between and within sectors," *Economics of Innovation and New Technology*, 12(6), 481–505.

VEUGELERS, R., and B. CASSIMAN (2005): "R&D cooperation between firms and universities. Some empirical evidence from Belgian manufacturing," *International Journal of Industrial Organization*, 23(5-6), 355–379.

VIVARELLI, M. (2012): “Innovation, Employment and Skills in Advanced and Developing Countries: A Survey of the Literature,” *IZA Discussion Papers* 6291, Institute for the Study of Labor (IZA).

VIVARELLI, M., and D. AUDRETSCH (1998): “The Link between the Entry Decision and Post-Entry Performance: Evidence from Italy,” *Industrial and Corporate Change*, 7(3), 485–500.

WOOLDRIDGE, J. M. (2008): *Econometric Analysis of Cross Section and Panel Data, 2nd Edition*, vol. 1 of *MIT Press Books*. The MIT Press.

WOOLDRIDGE, J. M. (2015): “Control Function Methods in Applied Econometrics,” *Journal of Hu- man Resources*, 50(2), 420–445.

ZETLIN-JONES, A., and A. SHOURIDEH (2012): “External Financing and the Role of Finan- cial Frictions over the Business Cycle: Measurement and Theory,” 2012 Meeting Papers 321, Society for Economic Dynamics.

ZINGALES, L. (1998): “Survival of the Fittest or the Fattest? Exit and Financing in the Trucking Industry,” *Journal of Finance*, 53(3), 905–938.