TITLE OF DISSERTATION: ESSAYS ON APPLIED INDUSTRIAL ORGANIZATION IN PAKISTAN

DISSERTATION
SUBMITTIED TO THE DEPARTMENT OF ECONOMICS
AND THE COMMITTEE FOR ADVANCED STUDIES AND RESEARCH
OF LAHORE SCHOOL OF ECONOMICS
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Maryiam Haroon
August 2021

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality to be considered for the degree of Doctor of Philosophy.

(Dr. Azam Chaudhry) Principal Supervisor

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality to be considered for the degree of Doctor of Philosophy.

(Dr. Theresa Chaudhry) Advisory Committee Member

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality to be considered for the degree of Doctor of Philosophy.

(Dr. Waqar Wadho) Advisory Committee Member

TITLE OF DISSERTATION: ESSAYS ON APPLIED INDUSTRIAL ORGANIZATION

by

Maryiam Haroon

Submitted to the Department of Economics on August, 2021 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Economics

Abstract

The recent literature has documented surprising differences in firm-level performance within countries and between developing and developed countries. The recent literature looking at developed economies highlights cluster policies and international trade shocks as potential mechanisms that boost productivity and quality upgrading of a firm's products. There is limited evidence in the literature examining these from a developing country perspective where cluster policies and international trade shocks.

In this thesis, we use a firm-level data set to answer three sets of questions: (i)do agglomeration economies in the form of localization, urbanization, and competitionincreasethe productivity and markups of firms; (ii) does the bilateral trade agreement between Pakistan and China enable firms to change their product mix and quality of their products;(iii) do agglomeration economies in the form of localization and urbanization economies attract new firms to locate in an area. Our analysis focused on firms from the province of Punjab in Pakistan.

We do not find evidence of agglomeration externalities improving the productivity of firms though we do find that firms in agglomerated regions have higher costs, and these firms also charge higher prices and increase markups. We also find that due to the free trade agreement Pakistani firms added more products because of lower Chinese tariffs. Additionally, firms improved product quality, lowered prices, and increased the quality of their inputs as a result of lower Pakistan tariffs on imported Chinese inputs. Finally, we find that firms are attracted to areas where other firms are engaged in similar activities.

Thesis supervisor: Dr. Azam Chaudhry
Supervisor's Title: Professor and Dean of the Faculty of Economics at the Lahore School of Economics

Acknowledgements

I owe a debt of gratitude to my advisor, Dr. Azam Chaudhry, for his support and guidance during the Ph.D. process. He has been generous in sharing his time and knowledge while also showing continued patience to allow me to work at my own pace in a self-directed manner.

I was very fortunate to have Dr. Naved Hamid as my mentor. I am deeply grateful for his guidance, support, invaluable advice, and mentorship. He showed confidence in my ability and encouraged me to take initiatives, pushed me beyond boundaries, and always celebrated my achievements.

Likewise, I am grateful to thank Dr. Shahid Amjad Chaudhry, who has provided a supportive and collaborative environment for research. I would also like to thank my supervisory team for their feedback on my work.

I have been truly blessed to have a great support system in the form of friends and family. I would like to thank Zehra Gardezi, and Farah Said for their ready help whenever I needed it and their sincere advice. I thank Zenab Naseem, Zunia Saif, and other CREB colleagues who listened to my rants, suggested solutions, and made some of the challenging times more manageable. I am grateful to my mother for her prayers and understanding during this long process and my sisters and niece for their support even when they had no idea what I was doing. I would especially like to thank Minaa Haroon for her constant support, encouragement and for helping me find solutions to every problem.

I dedicate this dissertation to my late father.

Table of Contents

List of Tables

List of Figures

1. Introduction

Firm performance has direct implications for the overall growth of an economy. Higher productivity potentially allows firms to produce a higher level of output with the same level of inputs and earn higher revenues which in turn can lead to higher GDP. However, the recent literature has documented surprising differences in firm-level performance within countries and between developing and developed countries (Syverson, 2011; Hsieh and Klenow, 2009). The productivity gap between firms in developed and developing countries has found to be especially substantial and the recent literature has focused on both the obstacles and opportunities in bridging this gap. Evidence shows that firms are much less productive in developing economies than in advanced economies with as some of the research has found that (according to 2014 data) productivity (as measured by total factor productivity) in the US was 1.7 times larger than in Mexico, and 2.6 times larger than in India (Feenstra et al., 2015). Firms are also constrained in terms of their ability to innovate, exploit new technologies, move toward producing more sophisticated products, diversity their product mix and react to the local and international. The literature highlights several obstacles facing firms such as weak institutions, lack of infrastructure, limited access to finance, bad management practices, difficulty in technology adoption, infrastructure, international supply chains, and misallocation of resources from less productive to more productive firms (Syverson, 2011) and many of these can be more severe in the case of developing countries.

Pakistan's manufacturing sector growth has stagnated over the last six decades (1960-2012) with manufacturing growth averaging approximately 6.3 percent over this time. However, over the same period, many economies in the same region grew more rapidly, an example of which is China which being experienced GDP growth of 9.1 percent between 1970-2009 (CIA World Fact Book). While the research recognizes that growth manufacturing sector can have a multiplier effect on overall growth and employment there has been little in the way of a coherent strategy besides discussions on the need for firms to move up the value chain (Afraz, Hussain, and Khan, 2014).

One policy that has been aggressively pursued by the Pakistani government over the last two decades is the development of industrial zones. Policy makers have seen industrial zones as a way for agglomeration economies to spur the entrance of new firms, improve firm productivity and improve the quality of their products. The analysis of agglomeration economies is especially interesting in the case of Pakistan since the country is characterized by naturally occurring industrial clusters which vary significantly in age (from a few decades old to almost a century old) and the number and nature of firm clusters are growing over time. Also, since the country faces some of the same problems as many developing economies, such as weak formal institutions, less efficient markets, costly research and development, weak contract enforcement, and binding financial constraints, agglomeration economies may not only lead to firm growth but may also lead to higher productivity, higher markups, reduce costs, and attract new firms. In first past of the thesis, we aim to understand if agglomeration economies are crucial for firms in developing country context. Our study has important implications for economic development and public policy in the context of a developing country in particular, and for industries characterized as low-skilled or less technologically advanced in general. This also has implications for government policy aimed at countering regional disparity in developing countries. We believe that this has promising implications for industrial policy aimed at building industrial zones and special economic zones in developing countries[2].

There has also been a policy focus in Pakistan on engaging in bilateral trade agreements to improve firm performance in terms of their productivity and quality of products to help firms climb the quality ladder and shift focus from low value-added goods to higher value-added goods. Pakistan and China signed a bilateral trade agreement to reduce tariffs on both sides. In the second part of the thesis, we study how a developing country like Pakistan is affected by a trade agreement with China.As China continues to cultivate economic ties with developing countries, it is important to gauge the economic impact of wide-ranging trade agreements on firms in these countries.

The research in this dissertation aims to understand firm dynamics by focusing on three core questions: (i) Does spatial proximity plays a role in making firm's more productive and generating higher markups[3]? What are the potential channels which are driving these? Is there heterogeneity in our findings between newer or older clusters? (ii) Do tariff reductions enable firms to change their product mix and move to higher quality products? Which particular type of tariff reductions derive these results? (iii) What are the factors that attract firms to locate in specific areas?

Industrial clusters are generally regarded as good for the productivity, growth, and development of a region for both rich and poor countries. The conventional economic wisdom dates back to Marshall's (1902)'s influential work which viewed industrial

---

[2] However, it could be interesting to determine how naturally occurring clusters develop and there could be potentially different benefits from government developed clusters to those of natural ones.

[3] The agglomeration literature cites greater competition as a potential feature of cluster, which would lead to lower mark-ups but we are trying to understand if collusion or other benefits of agglomeration increase them.

clusters as productivity-enhancing through pro-competitive pressures they may foster. According to the New Economic Geography literature, localization externalities (often referred to as Marshallian externalities) operate through input markets, labor markets, and knowledge spillovers. The externalities were considered to be a key factor in the success of Silicon Valley (Marshall, 1920). Localization externalities are believed to raise the productivity of firms directly, or indirectly through increasing profits by reducing costs or raising the price firms can charge for their products (Fafchamps and Said El Hamine, 2017). The diversity of industries within a region generates urbanization economies (Rosenthal and Strange, 2001; Combes, et. al., 2011) and inter-industry agglomeration economies connected with variety (Jacobs, 1984; Glaeser, et. al., 1992; Cainelli and Iacobucci, 2012) further emphasized by Jacobs (1984) who puts forward the argument that industrial diversity generates pecuniary externalities in the form of output and input linkages and transfer of ideas and technologies. Industrial clusters may indeed be cost-reducing and productivity enhancing, but there is an even older concern – dating back to at least Adam Smith – that gathering competitors in the same locale could instead lead to non-competitive behavior.

It may seem paradoxical that multiple producers in the same area would lead to non-competitive behavior rather than increased competition, but proximity facilitates easy communication and observation, which are theoretically (e.g., Green and Porter (1984), in the case of tacit collusion) and empirically (Brooks, Kaboski and Li, 2016; Marshall and Marx, 2012; Genesove and Mullin, 1998) associated with collusive behavior. Thus, one will not be surprised that spatial proximity can also support cooperative arrangements that would translate in reducing their costs or charging higher prices resulting in extracting higher markups by firms. There is extensive theoretical and empirical literature documenting both the positive and negative externalities arising from agglomeration economies (for a detailed review see: Rigby and Essletzbichler, 2002; Duranton and Kerr, 2015; Rigby and Brown, 2015; Duschl, et. al., 2015; Grillitsch and Nilsson, 2017).

Evaluation of naturally occurring agglomeration externalities and their effect on firm productivity and markups are particularly relevant to Pakistan's manufacturing firms as Pakistan is still at a stage of development where it can benefit from agglomeration economies. Using unique firm-level panel data, we contribute to the existing literature by evaluating the impact of natural agglomeration on total factor productivity, markups, costs, and prices while controlling for unobserved heterogeneity across firms for manufacturing firms in Punjab, Pakistan. In the first chapter of this thesis, we test multiple hypotheses: first, we evaluate how much a firm benefits in terms of productivity when other firms from the same sector or another sector are located nearby. Second, we test whether agglomeration gives firms a cost advantage. Third, we aim to understand how agglomeration affects a firm's ability to extract markups. Third, we try to understand the role of agglomeration externalities in determining prices and costs which could be the potential mechanism through which markups might be affected.

The literature linked to industry dynamics has considered the allocation of resources as an important topic relevant to a firm's entry and exit decision. However, it abstracts from the reallocation decisions of output made within multiproduct firms when firms adjust their products by adding or dropping products(Eckel and Neary, 2010; Bernard, et. al., 2010). There are substantial gains in aggregate output when policy reforms, such as international trade liberalization, or changes in market fundamentals induce a reallocation from low- to high-performance firms within industries. The recent literature adds a new dimension by incorporating reallocation of output within a multiproduct firm through changes in product mix in response to changes in the economic environment (Goldberg, et. al., 2010; Topalova & Khandelwal, 2010; De Loecker, et. al., 2016; Brandt, et. al., 2017; Copestake, 2020).There could be several possible mechanisms through which trade policy can affect firms. Due to increased import competition, domestic firms may act defensively by switching product if faced with low-cost imports or may focus on their 'core competencies' by dropping their least-productive products and increasing sales per product when facing increased competition from imports (Eckel & Neary, 2010). Increased import competition can pressure domestic firms into improving their efficiency (Holmes & Schmitz, 2001; Nishimizu & Robinson, 1984) and encourage them to exploit economies of scale. Import competition may further lead to a reallocation of resources if high-cost producers are forced to exit the market, which frees up resources for the efficient firms that survive (Roberts & Tybout, 1991). The other possible mechanism could be the creation of export opportunity which enable firms to move towards a few successful products and focusing on these selected products enables them to survive and grow or expanding their scope or changed their core products as an offensive strategy translating into higher productivity (Eckel and Neary, 2010; Iacovone and Javorcik, 2010). Lower input tariff increased access to cheaper imported raw material and greater competition in domestic input markets translates into substantial increases in product varieties and quality across firms as a result of cheaper or previously unavailable imported inputs (Goldberg, et. al., 2010; Topalova and

Khandelwal, 2011; Brandt et al., 2017; Bigsten, Gebreeyesus and Söderbom, 2016; Amiti & Konings, 2007).

The second paper of this thesis draws on an existing extensive theoretical and empirical literature on multiproduct firms which builds on theories of industry dynamics by modeling endogenous product selection of firms (Bernard, Redding and Schott, 2007).We use a newly constructed panel dataset for firms in Pakistan, drawing on data from the Census of Manufacturing Industries (CMI) from before and after the implementation of the Free Trade Agreement (FTA) between Pakistan and China. Our research adds to a thin but growing literature that disentangles the quality and productivity gains that arise from reductions on tariffs on final and intermediate inputs due to bilateral trade agreement (Amiti & Konings, 2007; Brandt et al., 2017; Copestake, 2020; Lovo & Varela, 2020)—both important channels through which China's integration with the global trading system has affected firms in developing countries. Additionally, by looking at the trading relationship between China and a developing country, this paper analyses the potential benefits of cheaper and higher-quality Chinese inputs for industrial upgrading, which is broadly applicable to many developing countries. Specifically, we test (a) the impact of this trade agreement on firm-level outcomes, and (b) disentangle the impact of greater access to export markets (due to lower foreign tariffs on exports) from the impact of greater competition in domestic markets (due to lower domestic tariffs on imported products) and greater competition in input markets (due to lower domestic tariffs on imported raw materials).

The concentration of industrial activity occurs widely across economies with well-known examples from developed economies, like the computer industry in Silicon

Valley (Sorenson and Audia, 2000), and in developing economies, like the surgical goods and sports goods industries in Sialkot, Pakistan (Atkin et al., 2015a; Nadvi, 1999). The industrial organization literature highlights agglomeration as one of the main factors attracting new firms in an area which is supported by the empirical literature. The empirical literature documents benefit from localization economies in the form of input sharing (because of accessibility to suppliers and mutually enforced contracts), labor pooling (because of the availability of specialized labor), and knowledge spillovers (because of shared information about products, production process, innovations, existing and new technology, marketing agendas, and research and development) (see Parr, 2002; Marshall, 1920). The literature also documents the benefit of urbanization economies as the presence of diversified suppliers, specialized labor and suppliers, market mechanisms, transportation facilities, infrastructure, and community facilities, which make certain areas more attractive for new firms to enter (Parr, 2002). Along the same lines, Sorenson and Audia (2000) found that new entrepreneurial activity is likely to take place in areas of geographic concentration while literature also document that new firms enter when they anticipate a developed market, existing suppliers, and the availability of low-cost factors of production. A common thread running through much of the empirical literature is that most of the research studying where firms decide to locate has focused on developed economies (Glaeser and Kerr, 2009; and Rosenthal and Strange, 2010). Regional economic characteristics and governmental policies have also been found to make a difference in the location decision of firms. Otsuka (2008) found that various location factors significantly affect the formation of new establishments in a particular region, including market demand, agglomeration, market conditions, and factor cost.

Reynolds et al. (1994) found that government policies attract new firms to a particular area through government spending on local infrastructure and the provision of direct assistance to firms.

In the third chapter of this thesis, we empirically estimate the relationship between agglomeration forces i.e., localization and urbanization economies, and the formation of new firms using industry district-level data from Punjab, Pakistan. We test two hypotheses: First, do new firms in an industry choose to locate in an area where there is a similar industrial activity? Second, do new firms chose to locate in an area where there is diverse industrial activity? We do this by analyzing the effects of agglomeration on the arrival and scale of operations of new firms at the district level in 2010, incorporating industrial controls and socioeconomic characteristics at the district level using a combination of firm-level data and household survey data.

References

Amiti, M., & Konings, J. (2007). Trade liberalization, intermediate inputs, and productivity: Evidence from Indonesia. *American Economic Review*, *97*(5), 1611–1638.

Afraz, N., Hussain, S. T., & Khan, U. (2014). Barriers to the growth of small firms in Pakistan: A qualitative assessment of selected light engineering industries. The Lahore Journal of Economics, 19, 135.

Atkin, D., Chaudhry, A., Chaudry, S., Khandelwal, A., & Verhoogen, E., (2015b). *Organizational barriers to technology adoption: evidence from soccer-ball producers in Pakistan.* NBER Working Paper No. 21417. doi: 10.3386/w21417.

Bigsten, A., Gebreeyesus, M., & Söderbom, M. (2016). Tariffs and firm performance in Ethiopia. *Journal of Development Studies*, *52*(7), 986–1001.

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2007). Firms in international trade. *Journal of Economic Perspectives*, *21*(3), 105–130.

Bernard, A. B., Redding, S. J., & Schott, P. K. (2010). Multiple-product firms and product switching. *American Economic Review*, *100*(1), 70–97.

Brandt, L., Van Biesebroeck, J., Wang, L., & Zhang, Y. (2017). WTO accession and performance of Chinese manufacturing firms. *American Economic Review*, *107*(9), 2784–2820.

Brooks, W. J., Kaboski, J. P., & Li, Y. A. (2016). *Growth Policy, Agglomeration, and (the Lack of) Competition* (No. w22947). NBER Working Paper Series. National Bureau of Economic Research.

Cainelli, G., & Iacobucci, D. (2012). Agglomeration, Related Variety, and Vertical Integration. *Economic Geography, 88* (3), 255-277.

Central Intelligence Agency. (2009). *The CIA world factbook 2010.* Skyhorse Publishing Inc..

Copestake, A. (2020). *Multi-product firms, networks and quality-upgrading: Evidence from China in India* (Working Paper No. 1659). San Rafael, CA: Forum for Research in Empirical International Trade.

Combes, P. P., Duranton, G., & Gobillon, L. (2011). The Identification of Agglomeration Economies. *Journal of Economic Geography, 11* (2), 253-266.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., & Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica, 84*(2), 445–510.

Duranton, G., & Kerr, W. R. (2015). *The Logic of Agglomeration* (No. W21452). National Bureau of Economic Research

Duschl, M., Scholl, T., Brenner, T., Luxen, D., & Raschke, F. (2015). Industry-Specific Firm Growth and Agglomeration. *Regional Studies, 49*(11), 1822-1839.

Eckel, C., & Neary, J. P. (2010). Multi-product firms and flexible manufacturing in the global economy. *The Review of Economic Studies, 77*(1), 188–217.

Fafchamps, M., & El Hamine, S. (2017). Firm Productivity, Wages, and Agglomeration Externalities. *Research in Economics*, 71(2), 291-305.

Feenstra, R, R Inklaar and M Timmer (2015), The next generation of the Penn World Table. *American Economic Review*, 105(10), 3150-3182.

Genesove, David, and Wallace Mullin. 1998. Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914. *RAND Journal of Economics*, 29(2): 355–377.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in Cities. *Journalof Political Economy*, *100* (6), 1126-1152.

Glaeser, E. L., & Kerr, W. R., (2009). Local industrial conditions and entrepreneurship: How much of the spatial distribution can we explain?. *Journal of Economics & Management Strategy*, *18*(3), 623-663. doi: 10.1111/j.1530-9134.2009.00225.x

Green, E. J., & Porter, R. H. (1984). Noncooperative Collusion Under Imperfect Price Information. *Econometrica: Journal of the Econometric Society*, 87-100.

Grillitsch, M., & Nilsson, M. (2017). Firm Performance in The Periphery: On the Relation Between Firm-Internal Knowledge and Local Knowledge Spillovers. *Regional Studies*, *51*(8), 1219-1231.

Goldberg, P. K., Khandelwal, A. K., Pavcnik, N., & Topalova, P. (2010). Multiproduct firms and product turnover in the developing world: Evidence from India. *The Review of Economics and Statistics*, *92*(4), 1042–1049.

Hsieh, C. T., & Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly journal of economics*, *124*(4), 1403-1448.

Holmes, T. J., & Schmitz, J. A., Jr. (2001). A gain from trade: From unproductive to productive entrepreneurship. *Journal of Monetary Economics*, *47*(2), 417–446.

Iacovone, L., & Javorcik, B. S. (2010). Multi□product exporters: Product churning, uncertainty and export discoveries. *The Economic Journal*, *120*(544), 481–499.

Jacobs, J. (1984), *Cities and the Wealth of Nations*, Random House, New York.

Lovo, S., & Varela, G. (2020). Internationally Linked Firms, Integration Reforms and Productivity.

Marshall, A. (1920). *Principles of Economics: An Introductory Volume*. Royal Economic Society(Great Britain).

Marshall, Robert, and Leslie Marx. (2012). The Economics of Collusion. MIT Press.

Nadvi, K. (1999). Shifting ties: Social networks in the surgical instrument cluster of Sialkot, Pakistan. *Development and Change,* 30 (1), 141-175. doi: 10.1111/1467-7660.00110

Nishimizu, M., & Robinson, S. (1984). Trade policies and productivity change in semi industrialized countries. *Journal of Development Economics*, *16*(1–2), 177–206.

Otsuka, A. (2008). Determinants of new firm formation in Japan: A comparison of the manufacturing and service sectors. *Economics Bulletin*, *18*(4), 1-7. File URL:

http://economicsbulletin.vanderbilt.edu/2008/volume18/EB08R30002A.pdf.

Parr, B. J., (2002). Missing elements in the analysis of agglomeration economies. *International Regional Science Review*, 25(2), 151–168. doi: 10.1177/016001702762481221.

Reynolds, P., Storey, D. J., & Westhead, P. (1994). Cross-national comparisons of the variation in new firm formation rates. *Regional Studies*, 28(4), 443–456. doi:10.1080/00343409412331348386.

Rosenthal, S. S., & Strange, W. C. (2001). The Determinants of Agglomeration. *Journal of Urban Economics*, *50*(2), 191-229.

Rosenthal, S. S., & Strange, W. C. (2010). Small establishments/big effects: Agglomeration, industrial organization, and entrepreneurship. *Agglomeration Economies*, The University of Chicago Press, (pp.277 – 302).

Roberts, M. J., & Tybout, J. R. (1991). Size rationalization and trade exposure in developing countries. In R. E. Baldwin (Ed.), *Empirical studies of commercial policy* (pp. 169–200). Chicago, IL: University of Chicago Press.

Rigby, D. L., & Essletzbichler, J. (2002). Agglomeration Economies and Productivity Differences in Us Cities. *Journal of Economic Geography*, *2*(4), 407-432.

Rigby, D. L., & Brown, W. M. (2015). Who Benefits from Agglomeration? *Regional Studies*, *49*(1), 28-43.

Syverson, C. (2011). What determines productivity? *Journal of Economic literature*, *49*(2), 326 -65.

Sorenson, O., & Audia, P. G. (2000). The social structure of entrepreneurial activity: Geographic concentration of footwear production in the United States, 1940-19891. *American Journal of Sociology*, *106*(2), 424-462. doi: 10.1086/3169.

Topalova, P., & Khandelwal, A. (2011). Trade liberalization and firm productivity: The case of India. *The Review of Economics and Statistics*, *93*(3), 995–1009.

2. Paper I: Measuring the Impact of Agglomeration Economies on Naturally Rising and Declining Manufacturing Sectors in Pakistan

Section 2.1: Introduction

Economic theory has focused on the impact of industrial clusters on productivity and growth for both developed and developing countries. The conventional economics wisdom dates back to Marshall's (1920) view of industrial clusters as productivity-enhancing because of the pro-competitive pressures that they foster, and more recent economic thought has also proposed that clusters can have a positive impact. At the same time, some of the literature has focused on how clusters can foster non-competitive behavior (Brooks, Kaboski, and Li, 2016). In this paper, we explore productivity, costs, and markups, i.e. efficiencies and competition in agglomerated areas, and whether these are different for younger vs. older clusters.

The theoretical literature is well developed and has identified positive (for example, shared infrastructure) and negative (for example, congestion) externalities (Fafchamps and Said El Hamine, 2017). However, empirical evidence on the precise nature of agglomeration externalities remains inconclusive. This paper draws an attention towards these externalities and its role in productivity, costs and markups of firms. The New Economic Geography literature highlighted the role oflocalization economies (often referred to as Marshallian externalities) and urbanization economies in determining productivity, costs and markups. The literature highlights that these externalities can raise the firm's productivity directly; or indirectly through increasing profits by reducing costs or raising the price of firms for their products (Fafchamps and Said El Hamine, 2017). Localization economies generate three sources of economies through which are labor pooling, input sharing and knowledge spillovers which are predicted to be facilitated in concentrated regions and raises productivity and could lead to cost benefits. Spatial proximity also helps in informational spillovers, in particular, in the form of business opportunities or market-relevant knowledge which can led to costs advantages, productivity and setting prices (Rauch and Casell, 2003; Fafchamps and Minten, 2002). Pecuniary externalities have also been proposed as possible explanation for spatial concentration (Fujita, Krugman and Venables, 1999) and improving productivity or reducing costs for firms, for instance, in thick or large labor markets it is easier for firms to find specialized workers (Glaeser et al. 1992).

In addition to localization externalities, the literature has also emphasized the importance of urbanization economies (Rosenthal and Strange, 2001; Combes, et., al., 2011) and inter-industry agglomeration economies connected with variety (Jacobs, 1984; Glaeser, et al. 1992; Cainelli and Iacobucci, 2012). Jacobs (1984) proposed that it is the diversity of industries within a region, which is the source of externalities since firms borrow ideas and technologies from each other. Industrial diversity generates pecuniary externalities such as output and input linkages directly impacting productivity. Forward and backward linkages is another proposed benefit in spatially concentrated regions which are crucial for firm development and functions Hirschman (1958). For instance, greater number of firms might foster entry of intermediate inputs in nearby areas thereby, generating specialization gains (Ciccone and Matsuyama, 1996).The empirical literature documenting both the positive and negative externalities arising from agglomeration economies presents mixed evidence (for a detailed review see: Rigby and Essletzbichler, 2002; Duranton and Kerr, 2015; Rigby and Brown, 2015; Capasso, Stam, and Cefis, 2015; Duschl, et al. 2015; Fontagné and Santoni, 2016; Grillitsch and Nilsson, 2017; Howell, et al. 2016).

The Schumpeterian literature has also discussed the impact of agglomeration. Most of this work has looked at how competition creates incentives for firms to invest and innovate, reducing costs and lowers prices (Porter, 1990). This literature has discussed how spatial proximity can foster competition, which can reduce the costs of inputs. While much of the literature looks at the positive impacts of agglomeration, another strand of the literature (dating back to Adam Smith) discusses how gathering firms in the same geographic area can lead to non-competitive behavior. Proximity facilitates communication and observation, which has been theoretically (e.g., Green and Porter (1984) and empirically (Brooks, Kaboski and Li, 2016; Marshall and Marx, 2012; Genesove and Mullin, 1998) shown to lead to collusive behavior. This collusive behavior can enable firms to pay less for inputs and set higher prices to extract higher markups.

This paper contributes to the existing literature in multiple ways: first, we construct a unique firm-level panel dataset for manufacturing firms in Punjab, Pakistan. This data also allows us to use actual product level prices to compute productivity and markups, which is largely not possible in the literature because of the unavailability of data. Second, our analysis adds to the relatively limited literature on the impact of agglomeration in developing economies; finally, in this paper, we explore a new dimension of agglomeration by looking at how agglomeration externalities can differ based on the cluster development time. Here we can exploit the fact that as developing countries grow there are two types of clusters: on one hand we have well-established clusters and on the other hand we have newer clusters that are in the process of developing. Using data from clusters in Punjab, Pakistan we look at the overall impact of agglomeration on productivity, costs, prices, and markups of firms in Punjab, Pakistan. We then decompose clusters into two types: firms from older clusters and firms from newer clusters and see if there is a differential impact of agglomeration on the same variables[4].

Our findings suggest that that localization has a negative correlation with the productivity of firms and these results are driven by the firms in older clusters. On the other hand, we do not find consistent evidence of urbanization economies with productivity. In terms of costs, prices and markups we find that urbanization increase both costs and prices of firms from newer clusters. These new cluster firms also benefit from higher marks up as opposed to older clusters whose cost prices and markups are unaffected by agglomeration. This seems to imply that the impact of agglomeration rather than being constant over time varies across sectors within a country that is in different stages of development.

Section 2.2: Why is Pakistan an interesting case?

While many researchers have analyzed the impact of agglomeration on firm-level outcomes in developed economies, there is significantly less literature on clusters in developing economies. What makes Pakistan an interesting case for a study of agglomeration externalities is that it is characterized by naturally occurring industrial clusters which vary significantly in age (from a few decades old to almost a century old) and the number and nature of firm clusters are growing over time.

One reason for this continued growth in clusters is that there is still significant scope for firm growth in Pakistan. The overall agglomeration index (based on urbanization) suggests that

[4] A detailed review of existing literature provides a comparison of cluster at different stages of development and its impact (Basant, 2002).

the country has not yet reached a stage where congestion hinders firm growth (World Development Report, 2009). Also, there are still significant geographical areas that have less firm-level activity. Looking at the level of the geographical spread of activity in Punjab, Pakistan, one can observe that most of the firms are concentrated in the central part of the region (see Figure 2.1) and unequal distribution of activity across districts in Punjab with Central Punjab which includes Lahore, Faisalabad, Sialkot emerging as the main hub of activity in Punjab. Another interesting observation is that different industries exhibit different levels of concentration with some more spread across districts in Punjab and others exhibiting higher concentration (see Figure 2.2). Historically, some of the most concentrated industrial clusters are sports and athletics goods, scientific instruments, glass and glass products, pottery and china products, iron and steel industries, wearing apparel, and pharmaceutical industry whereas the least concentrated industries are footwear manufacturing, animal feed, non-ferrous metals, petroleum refining, petroleum products, and coal and beverage industry (Burki and Khan, 2010). Some industrial clusters in Pakistan such as Sialkot's surgical instruments cluster have deep local roots dating back to the 1800s (Nadvi, 1998).

But cluster growth in Pakistan is not only because of the geographic opportunities. Since the country faces some of the same problems as much developing economics, such as weak formal institutions, less efficient markets, costly research and development, weak contract enforcement, and binding financial constraints, agglomeration economies likely continue to play a key role in firm growth.

More readily available intermediate inputs or complementary factors to support production also create a more favorable environment for firms to operate. For small-scale firms, the adoption of new technologies tends to be difficult while undertaking research and development is costly, so those firms are very likely to benefit by locating in agglomerated areas. The literature on Pakistani clusters has also shown that firm proximity alternates for strong social and business networks (Nadvi, 1999) and that these networks enable firms to charge higher prices and benefit from lower costs.

Since there has been continual growth in the number of clusters, we find a situation where there are older clusters (such as the Surgical cluster in Sialkot that was established more than a century ago) as well as newer clusters (which have been established in the last few decades). The existence of older and newer clusters potentially allows us to see if the impact of agglomeration varies as clusters become older or if this impact stays relatively constant across different types of clusters.

Section 2.3: Data

In our analysis, we make use of two extensive firm-level datasets for Punjab, Pakistan to test the relationship between agglomeration externalities, markups, and productivity. Our primary data source is the Pakistan government's Census of Manufacturing Industries (CMI) for two time periods i.e., 2005-06 and 2010-11 for the Punjab province of Pakistan, which contains data on 3,528 manufacturing units, representing approximately 35–40% of firms in Punjab[5]. We focus on Punjab since it is the largest province in Pakistan population-wise and is responsible for nearly half of the manufacturing value-added in the country (Pakistan Bureau of Statistics). Also, the firm-level data from the other provinces are either unavailable or not representative. Using the Punjab CMI data, we create a unique two-period panel of firms. Since the data on these firms

[5] We used 2006 and 2011 datasets as the recent data of Census of Manufacturing Industries is not publicly available. The recent survey was conducted by a private agency and access to it for private researchers is not possible yet.

did not have common identifiers across years, we matched firms across the years. After matching the firms, we were able to create a balanced panel and an unbalanced panel of firms and for this paper, we utilized the balanced panel.

The dataset has information regarding firm-level output, prices, revenue, fixed assets, labor, wages, raw material usage, the value of raw materials, capital, investments, number and types of products produced, quantity and value of each product produced, ownership structure, firm location, city, province, and district. The advantage of this data is that we can observe actual the price of each product a firm produces. This allows us to calculate firm-level productivity and markups using panel-data based approaches that address many of the econometric concerns that exist in much of the empirical literature that looks at developing countries. We make use of firms having single plants and not having multiple plants.

Our empirical analysis also uses data from the Government of Punjab's Directory of Industries (DOI) for 2002, 2006, and 2011 to calculate various measures of agglomeration. The DOI is an extensive firm-level data set, which collects information on all firms in Punjab with more than 16,000 firms in each year, but the dataset does not contain detailed information on these firms. It has information on each firm's year of establishment, employment level, district, industry, and address. Using the addresses given in the data, we computed the longitude and latitude for each firm which allowed us to calculate distances between firms. Combining the two datasets enables us to provide deeper insights into the geographical characteristics of economic activity in Punjab.

We are also able to classify clusters in terms of older and newer industrial clusters. For this, we follow the methodology developed by Burki and Khan (2010) who classify industrial clusters into old and newer as old clusters as those who had high concentration levels in the pre-1990's period while classifying newer clusters as those clusters with low concentration levels in the pre-1990's period. One of the interesting features of these industrial clusters is that the older clusters are mostly more localized historically and have concentrated in a particular district. Some of the main sectors that make up the older clusters include sports and athletics goods, scientific instruments, glass and glass products, pottery and china products, iron and steel industries, wearing apparel, and pharmaceuticals. The newer clusters are primarily made of firms involved in footwear manufacturing, animal feed, ice factories, non-ferrous metals, petroleum refining, petroleum products, and coal and beverages. Authors like Burki and Khan (2010) and Atkin, et. al., (2017) also discuss how the newer cluster firms tend to more capital intensive, use more updated technology in their production processes, and produce higher quality goods.

Section 2.4: Methodology

We adapt a model similar to Martin, et al. (2011) in which they use a firm-level Cobb-Douglas production function of the form:

$$Q_{iyrt} = A_{iyrt} K_{iyrt}^{\alpha} L_{iyrt}^{\beta} \tag{2.1}$$

where $Q_{iyrt}$is the output of firm i of industry y in region r and time t, A_{iyrt} is the total factor productivity of firm i of industry y in region r and time t, K_{iyrt} is the capital stock of firm i of industry y in region r and time t and L_{iyrt} is the labor offirm i of industry y in region r and time t. We estimate the firm-level production function developed by Levinsohn and Petrin's (2003) and Wooldridge (2009) to calculate TFP using semi-parametric methods. To estimate total factor productivity, we used value-added, output, and revenues as an outcome to predict the TFP for each firm.

The authors assumed TFP of firm *i* is dependent upon local conditions or characteristics of a regionand firm level characteristics. The local conditions of an area have been incorporated using agglomeration externalities. Thus, the TFP function can be written as:

$$A_{iyrt} = (Agg_{yrt})\,(X_{iyrt}) \tag{2.2}$$

In this specification, TFP of a firm is a function of agglomeration in industry *y* region *r* and time *t* and firm-level characteristics represented as X_{iyrt} After a logarithmic transformation we can write the models specified in equations (2.2) as:

$$q_{iyrt} = \alpha k_{iyrt} + \beta l_{iyrt} + a_{iyrt} \tag{2.3}$$

$$a_{iyrt} = Agg_{iyrt} + x_{iyrt} + v_{iyrt} \tag{2.4}$$

We analyze the impact of agglomeration on the productivity of firms using a two-step estimation strategy. In the first step, we use equation (2.3) to estimate the firm-level production function used by Levinsohn and Petrin (2003) and Wooldridge (2009) to calculate TFP using semi-parametric methods. To estimate total factor productivity, we used the value-added, output, and revenues to predict the TFP for each firm. In the second step, we find the correlation of computed firm-level productivity with various agglomeration externalities and other firm-level controls[6]. In addition to analyzing the correlation of agglomeration with TFP, we will estimate the correlation with markups by replacing the dependent variables in equation (2.4). We

[6] Our estimations include firm-level controls including physical capital, ownership structure, and regional four sub-divisions of Punjab.

employed the methodology developed by De Loecker and Warzynski (2012) to estimate mark-ups[7]. This paper overcomes the problem of missing price data of output and uses actual prices to compute markups. From the computed input-output elasticity, we calculated markups.

We incorporate agglomeration in various forms. Following Combes, et. al., (2004), we incorporate the local environment through measures of localization (localization in industry *y* region *r* and time) and urbanization (urbanization in region *r* and time *t*) and competition (competition in industry *y* region *r* and time *t*). We will measure localization using the total size of a sector in a region (L_{yrt}) which a summation of employment in industry *y* region *r* and time *t*. Urbanization or diversity is measured in three ways: the total employment in a particular region, the total number of sectors present in a region, and the diversity index. The total employment in region (L_{rt}) is the summation of employment in region *r* and time *t*. The total number of sectors (S_{rt}) is the summation of the total sectors present in region *r*, and time *t*. The greater the number of sectors in a region, the more diversified it is. The diversity index (Div_{rt}) faced by each firm in region *r* and time *t* is defined as:

$$Div_{rt} = \frac{1}{\sum_{y \epsilon I_{rt}} \left(\frac{L_{yrt}}{L_{rt}}\right)^2} \qquad (2.5)$$

Lastly, competition is incorporated in our model by using the total number of firms in the sector and a competition index. The total number of firms in the sector *y* region *r* and time *t* is denoted by (N_{yrt}). The competition index (C_{yrt}) in sector *y* region *r* and time *t* is defined as:

7 We obtained the markup from the firm's cost minimization problem by substituting in estimates of the output elasticity of a variable input and the disturbance that separates actual from planned output.

$$C_{yrt} = \frac{1}{\sum_{x \epsilon I_{yrt}} \left(\frac{L_x}{L_{yrt}}\right)^2} \tag{2.6}$$

(where L_x is the employment of each firm x)

Our agglomeration measures are absolute and not relative measures. The literature does not consider one measure superior to another through one of the criticisms of absolute measures is that they do not account for area size. In light of such concerns, we also use a relative measure for localization and urbanization. We define relative measure for localization as location quotient {(employees in region r and industry y)/(total employees in industry y)} and relative measure for urbanization {(employees in region r)/(total employees inPunjab)}.

The recent literature also pointed out that the level of aggregation of data plays a crucial role in identifying the impact of agglomeration as these externalities are more often present as the level of geographical aggregation decreases (Beaudry and Schiffauerova, 2009). Since most of the literature has focused on analyzing agglomeration externalities at an aggregated level using provinces or districts as their unit of measurement, we expect that the agglomeration and productivity relationship is dependent upon the choice of aggregation of variables. Some of the recent research has adapted a distance-based methodology since this is likely to provide a more complete and unbiased analysis of the location patterns of industries (Lang, et. al., 2015). In line with this research, we calculate our agglomeration measures using 10, 5, and 2-kilometer radii which are used in our robustness tests.

The recent literature has also tried to move away from crude measures relying on economic density to measure agglomeration. With the development of economic geography literature, Duranton and Overman (2002) presented criteria for an index of spatial concentration which relies on a distance-based approach rather than selecting district states as their unit of measurement which is likely to suffer from statistical bias induced by the choice of spatial unit. Following this, we will also use the *'M function'* proposed by Lang, Macron, and Puech (2009). It can be calculated as the ratio of neighboring firms belonging to industry *'y'* within radius *'r'* over the number of all firms within that distance. This is then averaged across the industry and is compared to the ratio of all firms in the industry *'y'* over the total number of firms in Punjab. To control for the concentration within the plant and accounting for the size of each firm, we attach weights and use the number of employees rather than the number of firms. We calculated the M function using 10, 5, and 2-kilometer radii which we again use in our robustness tests.

There could be unobserved factors that can affect TFP across regions and firms, which might be correlated with our main agglomeration externalities. Using firm-level panel data allows us to address this issue by using firm-level fixed effects, which will take into account all firm-specific characteristics that are invariant across time. Our analysis will also control for unobserved heterogeneity using firms and year-fixed effects. Simultaneity bias can also be a problem because while higher agglomeration might make firms more productive, it is also possible that the presence of more productive firms in an area might cause other firms to locate nearby which increases agglomeration. In order to address this potential simultaneity bias, we instrument agglomeration measures with their lagged values. Lastly, we also aim to analyze the differential impact of agglomeration for the above-mentioned variables for firms in the older and newer clusters.

Section 2.5: Data Description

A preliminary analysis of the dataset (see Table 2.1a) shows some interesting facts about the distribution of firms. On average there are approximately 717 workers in each sector in a district with an increase from 2006 to 2011. Overall, there is a relatively high level of activity in each district with an average of 60,284 workers and 77 five-digit sectors operating in each district, though the index of diversity is relatively low. On average there are approximately 14 firms operating in each sector within a district in 2011 while the number of firms in each sector in a district in 2002 is much higher with an average of around 39 firms. We also observe that the competition index has decreased from 2002 to 2011. We look at the same statistics using 4 and 3 digit industrial classification and observe that as one moves to broader industrial classifications, the activity in each seems to be greater in quantity and lower in dispersion.

We also plot the total factor productivity of firms for the older and newer clusters in 2006 and 2011 in Figures 2.3 and 2.4. The plots of productivity distributions in 2006 show that the productivity of firms in the older and newer clusters are not significantly different in 2006. The productivity plots for 2011 show that firms in older clusters have slightly higher average productivity.

In table 2.1d, we look at the breakdown of firms in the older and newer clusters. Here we see that the number of newer clusters is almost three times the number of older clusters and the average number of firms per newer cluster is significantly higher than the average number of firms per older firm cluster.

Section 2.6: Findings and Analysis

We start with a standard analysis by testing the impact of agglomeration on productivity, markups, costs, and prices for all the firms. We then split the sample into firms from newer and older clusters and test to see the differential impact of each of these variables on productivity.

Subsection 2.6.1: Impact of agglomeration in the aggregate sample

i. Impact of agglomeration on firm productivity

Table 2.2 presents estimates of the correlation between total factor productivity and agglomeration. The estimations used TFP estimates from Levinsohn and Petrin (2003) and Wooldridge (2009).

Our findings suggest that when we look at all the firms together, the density of economic activity from the same sector is not found to have any effect on the productivity of firms which is mostly consistent across all specifications. These results contrast with much of the existing literature which found that the elasticity of productivity with respect to the size of the city or industry lies between 3% to 8% (Martin, Mayer, and Mayneris, 2011; Henderson, 2003; Rosenthal and Strange, 2003). This seems to imply that in the case of Pakistan, localization (Marshallian) externalities do not play an important role in improving the productivity of firms.

Surprisingly, we find that greater numbers of sectors present do improve firm-level productivity, which means that firm-level productivity improves if an area has more sectors operating nearby. One possible explanation for this could be that a greater number of sectors in an area might enable firms to benefit from diversified input and output suppliers which facilitates the transmission of ideas and technologies. Our findings are in line with what has been hypothesized by Jacobs (1984) and are in contrast to those found in the literature (Fafchamps and El Hamine, 2017; Glaeser, et. al., 1992; Henderson, 2003).

The presence of higher competition in closer proximity does not have an impact on firm-level productivity. This is in contrast to the results of Henderson (2003) who finds that greater competition has strong productivity effects, though Henderson's (2003) results were for high-tech and not machinery-based industries that make up the majority of the Pakistan manufacturing sector. One reason why we find no significant impact may be because most Pakistan industries tend to be characterized by low technology (relative to the international technology frontier) which does not spillover between firms. Or it could be that this variation is due to firm size as large firms tend to use higher technology than small firms. [Table 2.2 near here]

ii. Impact of agglomeration on markups

After testing the impact of agglomeration on productivity, we then test the impact on markups. Table 2.3 presents estimates of the correlations between mark-ups and agglomeration using a discrete space approach.

We find that firm's markups are increasing in localization externalities which suggests that if a firm is in closer proximity to other firms in the same sector, they are able to extract higher markups. The possible explanation for this could be because firms might be able to set higher prices collectively or pay lower costs since firms might be benefiting in sales, hiring, and input purchasing or price setting. Higher markups hint towards the possibility of collusion between firms since geographical proximity might enable easy cooperation and coordination thus enabling them to act jointly and exercise greater power. Our findings are robust to relative measures of agglomeration externalities. We do not find any impact of urbanization or competition externalities on firm-level markups [Table 2.3 near here].

iii. Are higher markups due to higher prices or lower costs?

In order to understand the mechanism through which firms might be able to extract higher markups in more agglomerated regions, we test whether higher markups are through cost advantage in agglomerated regions or is it through higher prices. Table 2.4 presents an estimate of the relationship between agglomeration and costs and Table 2.5 presents estimates for price and agglomeration.

In Table 2.4, we present estimates for three different types of costs: labor costs (column 2), raw material costs (column 3), and total costs (column 1). We find that the greater is the density of the same sector activity in an area, the higher are total and raw material costs (but not labor costs). One of the interpretations could be that high-quality goods require high-quality inputs which suggests that high cost in agglomerated regions reflects the use of high-quality inputs resulting in producing high-quality output (Kugler and Verhoogen, 2012; Atkin et. al., 2014). Also, greater competition among firms or a greater number of firms in an area results in lower total and raw material costs. This could be because a greater number of firms in the industry enable raw material supplying firms to enjoy economies of scale. The greater density of activity from other sectors and the more the diversified an area is in terms of the number of sectors, the lower is the total and raw material costs. [Table 2.4 near here]

Table 2.5 presents estimates for the relationship between agglomeration and the prices of the product. We use the weighted average price of products that a firm produces and the price of the highest revenue-generating product[8]. We don't find any significant relationship between the weighted average product price and any of our agglomeration measures (columns 1 and 2). We also find that the greater density of a sector enables a firm to charge higher prices for its highest revenue-generating product. This is in line with literature on collusive pricing behavior in

[8] We also estimated our model using quality adjusted prices and results are similar to the ones found for weighted average prices.

agglomerated regions (Brooks et al. 2017) which suggests that firms in closer proximity are more likely to collude that is firm charge higher markups under cartel decision than when firms operate independently. [Table 2.5 near here]

This suggests that the benefits firms were deriving in extracting greater markups are coming from cost-benefit and part of it is coming from higher prices.

Subsection 2.6.2: Differential impact of agglomeration across older and newer clusters

The analysis above follows the standard methodology for estimating the impact of agglomeration on firm-level variables like productivity, markets, prices, and costs. When we look at all of the firms together, the results imply that localization has a negative impact on productivity and that localization leads to higher prices and costs. In this section, we divided the firms into firms from newer clusters and firms from older clusters to see if there is a differential impact of agglomeration. Table 2.6 presents decomposed estimates for the relationship between agglomeration measures and productivity, markups, raw material costs, labor costs, and prices separately for older and newer clusters. [Table 2.6 near here]

i. Impact of agglomeration on firm productivity

Our initial results suggested that localization has a negative relationship with total factor productivity and estimates from table 2.6 (column 1) suggest that these results are being driven by firms in older clusters. Or in other words, localization has a negative impact on the productivity of firms from older clusters, but this impact does not exist for firms in newer clusters. This result may be due to the fact that older cluster might be using older technology or traditional practices which makes them unable to benefit from economies of scale (Basant, 2002).

On the other hand, our results show that firms in the newer clusters have lower productivity if the total employment in an area is higher (column 2). One of the possible reasons could be that newer cluster firms tend to employ more labor and also tend to use higher technology; so it is possible that the higher the level of total employment in an area, the more difficult it is for newer cluster firms to attract and keep better-trained labour which in turn impacts productivity[9]. At the same time, we find that firms from newer clusters have higher productivity if there are a greater number of sectors present in an area. In this case, it is possible that the greater number of total sectors reflects a higher level of economic activity in an area which allows firms in the newer cluster to benefit from economies of scale.

ii. Impact of agglomeration on markups

Our results also show that the greater the number of same sector firms in an area (represented by localization) then the higher the markups for firms in newer clusters (column 4). One possible explanation for this could be that the firms in the newer clusters tend to produce higher quality output which may provide them with greater market power. We also find that firms from newer clusters can extract higher markups if there are more sectors nearby; this could be because the total number of sectors represents the size of the local market, and a larger local market enables the newer cluster firms to extract higher markups.

iii. Are higher markups due to higher prices or lower costs?

The disaggregated results also show that firms in newer clusters have to pay higher raw

[9] These are the possible explanations of our results; however, we are not formally testing for these.

material and labor costs in the presence of greater localization economies which indicates diseconomies from localization (columns 6 and 8). This is consistent with our aggregate estimates and shows that the aggregate results are driven by firms in newer clusters as opposed to older clusters. This result makes sense since it is very possible that the firms from newer cluster have to compete with each other for higher quality raw materials which drive up their costs.

Another interesting finding is that firms in newer clusters have higher costs but they are able to charge higher prices as a result of localization. This again seems to be a reflection of the possibility that firms from newer cluster produce higher quality products are able to gain some market power which enables them to charge higher prices.

On the whole, our results seem to imply that the impact of agglomeration rather than being constant across different types of clusters tends to vary across clusters depending on the development, stage, and type of cluster[10].

Subsection 2.6.3: Robustness and additional checks

i. Impact of agglomeration on productivity

In this section, we present use an alternative technique to test the robustness of our results. Table 2.7 presents the results with agglomeration measured by m-function and Table 2.8 presents the decile estimates for the agglomeration and productivity relationship.

a. Robustness checks for total factor productivity and agglomeration

[10]Since we do not have the entire history or profile of these cluster it is difficult to attribute the effects to the stage of cluster alone and there could be other factors such as cluster or sector-specific factors, the age profile of the firms, the infrastructure in the clusters, the connections to road networks which could also potentially affect our main outcomes.

We estimate our model using different specifications to check the validity of our results. Our robustness includes estimations using different measures of agglomeration which includes relative measures of localization and urbanization (presented in Table 2.2 column 2 and 4) and an agglomeration index measured through the m function (presented in Table 2.7) and disaggregated analysis for newer and older clusters (presented in Table 2.7a). The findings also remain consistent when we employed alternative measurement techniques for measuring productivity (presented in Table 2.2 columns 3 and 4). They are also robust to using the data for private firms only and for single product firms only.

b. Do agglomeration externalities differentially impact the productivity of low, medium and high productivity firms?

In order to understand the importance of agglomeration economies for firms in different ranges of productivity distribution, we then examine whether it is low, medium, or high productivity firms that benefit most from agglomeration externalities. Since our initial estimations did not provide any conclusive results for agglomeration economies thus, we looked at the type of firms that are likely to benefit from these economies. Table 2.8 reports regression estimates from decile regressions where the deciles are based upon total factor productivity. We used an agglomeration index based upon m-function for these estimations for 5 kilometers radius. Our findings suggest that agglomeration economies are not beneficial for the total factor productivity of firms. We could not find any significant correlation of agglomeration economies with total factor productivity for firms lying in any of the firm's product range. In addition to analyzing the agglomeration economies, we analyzed whether localization economies have a role to play for firms lying in different ranges of the productivity distribution. We find that there are benefits of localization to firms lying in the medium range of the productivity distribution. The highest productivity firms are not affected by greater total activity and this may be because these productive firms are more focused on internal processes and are less affected by other firms. This provides us with the evidence that the presence of local industry is beneficial for some firms in the medium range of productivity distribution [Table 2.8 near here].

ii. *Impact of agglomeration on markups*

Table 2.9 presents regression estimates for the agglomeration and markups relationship using a continuous space approach with agglomeration variables measured within 10 and 5-kilometer radii rather than at the district level. Table 2.10 presents the results with the agglomeration measured by the m-function and Table 2.11 presents the decile estimates for the agglomeration and markups relationship.

a. *Do agglomeration externalities operate in narrowly defined geographical areas?*

When we consider agglomeration externalities in closer proximity using 10 and 5-kilometer radii, we find the magnitude of localization economies has increased and the strength of the localization economies has increased as well. This suggests that localization economies are more local and greater proximity of the sector enables firms to set or extract higher markups. However, we also find that greater competition between firms decreases markups. This suggests that greater competition decreases the ability of firms to act collusively or extract higher markups. These results are robust to both 10 and 5-kilometer radii [Table 2.9 near here].

b. Robustness estimations of mark-ups and agglomeration

Finally, we estimate our model using different specifications to check the validity of our results. Our robustness estimations use different measures of agglomeration which include relative measures of localization and urbanization (presented in Table 2.3 column 4 and 8) and the m-function based agglomeration index (presented in Table 2.10). The disaggregated analysis for newer and older clusters using m-function is presented in Table 2.10a. The results do not significantly change, and the results are also robust when we only use data for private firms (Table 2.3 column 3, 4, 7, and 8) and for single product firms only (Table 2.3 column 2 and 6). Our estimations measuring agglomeration using an m-function also provide strong evidence of agglomeration economies (Table 2.10). Finally, we also use different ways of calculating TFP by using the methods developed by Levinshon and Petrin (2003), Wooldridge and Olley and Pakes (1996) [Table 2.10 near here].

c. Do agglomeration externalities differentially impact low, medium and high mark-ups firms?

We also explore whether there is heterogeneity in the benefits that firms derive from agglomeration externalities according to the distribution of their markups. Table 2.11 reports the regression estimates from deciles regressions where the deciles are based on markups. We used the m-function agglomeration index for these estimations for both 10- and 5-kilometers radii. Our findings suggest that markups are increasing for firms lying in the 20^{th} to 90^{th} deciles of the markup distribution. However, markups do not increase with spatial proximity for firms lying at the top of the markup distribution [Table 2.11 near here].

Section 2.7: Conclusion

The evaluation of agglomeration externalities has gained considerable attention in the academic literature on industrial concentration. The presence of agglomeration externalities has been used to justify cluster policies by national and local governments in both developed and developing countries. In Pakistan, industrial clusters and special economic zones are a key area of focus for industrial policymakers attempting to expand the industrial base and increase industrial competitiveness. There is significant evidence that industrial clusters reduce costs and increase productivity, but there is also evidence that agglomeration may result in firms charging higher markups. Up to now, much of the empirical evidence on the impact of agglomeration has focused on developed economies.

In this paper, we test the impact of agglomeration in a developing country context using newly constructed firm-level panel data for Pakistan. Our analysis adds to the relatively limited literature on the impact of agglomeration in developing economies and also adds a new dimension to the study of agglomeration by looking at how agglomeration externalities can differ based on whether the firm is from a newer or older cluster. Using data from clusters in Punjab, Pakistan we look at the overall impact of agglomeration on productivity, costs, prices, and markups of firms in Punjab, Pakistan. We then decompose clusters into two types, older and newer clusters, and see if there is a differential impact of agglomeration on the same variables.

We do not find consistent evidence of agglomeration externalities improving the productivity of firms though we do find that firms in agglomerated regions have higher costs and these firms are able to charge higher markups. When we separate the firms into those from newer and older clusters, we find that the productivity of firms in the older clusters is negatively affected by agglomeration while the markups and costs of firms in the newer clusters are positively affected by agglomeration. The results seem to imply that a more nuanced approach to understanding the impact of agglomeration is required especially in countries that are characterized by both older and newer clusters of firms.

Table 2.1a: Summary statistics of agglomeration variables

Agglomeration Measures	Year 2011					Year 2006					Year 2002				
	N	Mean	St. dev	Min	Max	N	Mean	St. dev	Min	Max	N	Mean	St. dev	Min	Max
5-digit industrial classification															
Total size of the sector (L_{ijt})	1,183	717.2	3,106	10	70,827	1151	733.2	3365	10	70593	1,542	477.9	2,146	10	50,607
Total employment in location (L_{it})	1,183	60,284	56,585	1,053	161,399	1151	57837	51617	1080	141266	1,542	38,489	35,251	1,803	128,722
Total number of firms in the sector (N_{ijt})	1,183	14.4	48.81	1	1,069	1151	15.06	54.6	1	1280	1,542	39.37	207.6	1	6,905
Total number of sectors present in the location (M_{it})	1,183	77.05	57.16	4	180	1151	71.73	53.21	4	173	1,542	71.84	43.87	13	148
Diversity index (D_{it})	1,183	9.129	7.272	1.608	21.82	1151	7.333	5.747	1.337	21.6	1,542	9.713	7.784	1.883	30.47
Competition Index (C_{ijt})	1,183	7.341	18.05	1	327.3	1151	7.788	20.08	1	351.4	1,541	23.5	98.82	1	3,049
4-digit industrial classification															
Total size of the sector (L_{ijt})	909	933.4	3,605	10	70,827	890	948.2	4183	10	81341	1,217	605.5	2,458	10	50,662
Total number of firms in the sector (N_{ijt})	909	18.74	59.15	1	1,073	890	19.48	65.62	1	1284	1,217	49.88	234.5	1	6,909
Total number of sectors present in the location (M_{it})	909	45.36	29.15	4	102	890	43.92	27.83	4	98	1,217	48.8	25.42	12	99
Diversity index (D_{it})	909	7.395	5.792	1.608	19.47	890	6.016	4.76	1.337	18.56	1,217	8.415	6.547	1.882	27.88
Competition	909	8.848	20.46	1	298.1	890	9.372	24.1	1	404.4	1,216	24.43	104	1	3,051

Index (C_{ijt})															
	3-digit industrial classification														
Total size of the sector (L_{ijt})	697	1,217	4,841	10	84,943	689	1,225	5,259	10	84,882	942	782.2	3,145	10	51,757
Total number of firms in the sector (N_{ijt})	697	24.44	71.74	1	1,073	689	25.16	80.22	1	1,284	942	64.45	267.5	1	6,909
Total number of sectors present in the location (M_{it})	697	29.67	17.18	4	66	689	28.87	15.9	4	63	942	34.13	14.97	12	65
Diversity index (D_{it})	697	5.842	4.547	1.571	16.18	689	4.875	3.267	1.337	12.49	942	7.002	4.889	1.833	22.08
Competition Index (C_{ijt})	697	10.46	23.51	1	298.1	689	10.94	23.84	1	217.2	942	30.13	118	1	3,051

Note: Author's own calculation using Directory of Industries dataset for 2011, 2006 and 2002 using 5, 4 and 3-digit industrial classification to define sectors.

Table 2.1b: Summary statistics of firm characteristics

	2011		2006	
Variables	Average	St. dev	Average	St. dev
Output	24530	197330	34491	246544
Sales	1051922	5078395	1188333	63089
Capital	991542	5700043	216411	1097873
Labor	219	1067	175	527
Management Workers	160	555	146	459
Wages	38478	174090	19603	88432
Energy costs	66026	419192	30866	170593
TFP	7.07	1.45	6.97	2.02

Table 2.1c: Summary Statistics of firm characteristics across older and newer clusters

	Older cluster		Newer cluster	
Variables	Average	St. dev	Average	St. dev
Output	8417	50486	29775	225148
Sales	465384	2266781	1242852	5690301
Capital	260150	891543	1194963	6413090
Labor	161	459	236	1184
Management Workers	122	366	170	598
Wages	24566	63323	42437	194251
Energy costs	20572	77439	78940	472490
TFP	7.14	1.95	6.92	2.03

Table 2.1d: Summary statistics of older and newer clusters

	Older clusters	Newer clusters
Total number of clusters	33	88
Total number of firms	251	882
Average number of firms per cluster	7.6	10.02

Note: Author's own calculation using Census of Manufacturing Industries Dataset for 2011 and 2006

Table 2.2: Estimates of agglomeration and total factor productivity of firms (discrete space approach)

	TFP (Levinhson and Petrin)		TFP (Wooldridge)	
	(1)	(2)	(3)	(4)
Localization				
Total size of the sector	-0.156*		-0.156*	
	(0.086)		(0.0861)	
Relative localization		-0.111*		-0.111*
		(0.0667)		(0.0667)
Urbanization				
Total Employment	-0.966***		-0.966***	
	(0.359)		(0.359)	
Diversity Index	0.156		0.156	
	(0.262)		(0.262)	
Total Number of Sectors	0.717***		0.717***	
	(0.275)		(0.275)	
Relative Urbanization		0.138		0.138
		(0.141)		(0.141)
Competition				
Total number of firms in a sector	-0.0604		-0.0604	
	(0.165)		(0.165)	
Competition index	0.304		0.304	
	(0.206)		(0.206)	
Individual Firm Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Instrumented agglomeration	Yes	Yes	Yes	Yes
Level of Industrial classification	5 digits	5 digits	5 digits	5 digits
Level of Geographical	District	District	District	District

aggregation				
Types of firms	All	All	All	All

Note: All variables are in log form. Standard errors are in parenthesis. *** Indicate 1% significance level, ** indicates 5% significance level and * indicates 10% significance level. Agglomeration measures have been instrumented with past values using the previous census data.

Table 2.3: Estimates of agglomeration and markups of firms (discrete space approach)

	Markups (Levinhson and Petrin)				Markups (Wooldridge)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Localization								
Total size of the sector	0.014**	0.016*	0.016*		0.00882**	0.0102**	0.00914**	
	(0.00681)	(0.00897)	(0.00788)		(0.00343)	(0.00449)	(0.00393)	
Relative localization				0.138*				0.138*
				(0.0719)				(0.0719)
Urbanization								
Total Employment	-0.0433	-0.0291	-0.0442		0.00285	0.00724	-0.000856	
	(0.0283)	(0.0300)	(0.0318)		(0.0180)	(0.0191)	(0.0204)	
Diversity Index	-0.0230	-0.0387	-0.0313		-0.0172	-0.0220*	-0.0213*	
	(0.0209)	(0.0241)	(0.0245)		(0.0108)	(0.0122)	(0.0127)	
Total Number of Sectors	0.0524**	0.0408*	0.0560**		-0.00473	-0.0115	-0.00161	
	(0.0217)	(0.0244)	(0.0231)		(0.0146)	(0.0164)	(0.0157)	
Relative Urbanization				-0.150				-0.150
				(0.152)				(0.152)
Competition								
Total number of firms in a sector	-0.0194	-0.0202	-0.0187		-0.00489	-0.00402	-0.00545	
	(0.0132)	(0.0156)	(0.0152)		(0.00677)	(0.00777)	(0.00772)	
Competition index	0.00481	0.00955	0.0102		-0.00215	-0.00313	-0.000886	
	(0.0164)	(0.0172)	(0.0181)		(0.00920)	(0.00923)	(0.0102)	
Individual Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Instrumented agglomeration	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Level of Industrial classification	5 digits	5 digits	5 digits	5 digits	5 digits	5 digits	5 digits	5 digits
Level of Geographical aggregation	District	District	District	District	District	District	District	District
Single product firms	No	Yes	No	No	No	Yes	No	No
Types of firms	All	All	Private only	Private only	All	All	Private only	Private only

Note: All variables are in log form. Standard errors are in parenthesis. *** Indicate 1% significance level, ** indicates 5% significance level and * indicates 10% significance level. Agglomeration measures have been instrumented with past values using the previous census data. The outcome in all specifications is markups derived using TFP estimated through Levinhson and Petrin and Wooldridge.

Table 2.4: Estimates of agglomeration and costs

	Total costs	Labor costs	Raw material costs
	(1)	(2)	(3)
Localization			
Total size of the sector	0.300*	0.0246	0.275**
	(0.155)	(0.0510)	(0.135)
Urbanization			
Total Employment	-1.313**	0.0005	-1.314**
	(0.655)	(0.216)	(0.571)
Diversity Index	-0.847*	-0.126	-0.721*
	(0.472)	(0.155)	(0.411)
Total Number of Sectors	1.341***	-0.0323	1.374***
	(0.502)	(0.165)	(0.438)
Competition			
Total number of firms in a sector	-0.518*	-0.0093	-0.509*
	(0.300)	(0.0988)	(0.262)
Competition index	0.551	0.0422	0.508
	(0.376)	(0.124)	(0.328)
Individual Firm Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Instrumented agglomeration	Yes	Yes	Yes
Level of Industrial classification	5 digits	5 digits	5 digits
Level of Geographical aggregation	District	District	District

Note: All variables are in log form. Standard errors are in parenthesis. *** Indicate 1% significance level, ** indicates 5% significance level and * indicates 10% significance level. These estimation uses 5- digit industrial classification. The outcomes are total costs, labor costs and raw material costs.

Table 2.5: Estimates of agglomeration and price of product

	Weighted Average Price of all products		Price for Most profitable product	
	(1)	(2)	(3)	(4)
Localization				
Total size of the sector	0.300	0.372	3,747**	3,771**
	(0.361)	(0.363)	(1,579)	(1,672)
Urbanization				
Total Employment		-0.378		109.2
		(0.414)		(1,908)
Diversity Index		0.336		-2,288
		(0.340)		(1,567)
Total Number of Sectors		0.328		-1,504
		(0.434)		(2,001)
Competition				
Total number of firms in a sector	-0.509	-0.911	-5,497**	-5,115*
	(0.527)	(0.614)	(2,304)	(2,831)
Competition index		0.331		618.1
		(0.286)		(1,317)
Individual Firm Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Instrumented agglomeration	Yes	Yes	Yes	Yes
Level of Geographical aggregation	District	District	District	District

Note: All variables are in log form. Standard errors are in parenthesis. *** Indicate 1% significance level, ** indicates 5% significance level and * indicates 10% significance level. The outcomes are weighted average price of all products a firm produces and price of highest revenue generating product.

Table 2.6: Estimates of agglomeration and total factor productivity, markups, raw material costs, labor costs and price of product for older and newer clusters

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	TFP		Mark ups		Raw materials		Labor costs		Price of highest revenue generating product	
	Older cluster	Newer cluster	Older cluster	Newer cluster	Older cluster	Newer cluster	Older cluster	Newer cluster	Older cluster	Newer cluster
Localization										
Total size of the sector (L_{ijt})	-0.664*	-0.0823	-0.00472	0.0192**	-0.0609	0.270*	-0.0724	0.0979*	-58.02	3,740*
	(0.362)	(0.0985)	(0.0180)	(0.00811)	(0.274)	(0.159)	(0.237)	(0.0569)	(1,532)	(2,021)
Urbanization										
Total employment in location (L_{it})	-2.145	-1.096***	0.0299	-0.0528	-0.125	-1.219*	-0.557	0.0736	-384.9	-623.1
	(1.830)	(0.404)	(0.0914)	(0.0327)	(1.387)	(0.654)	(1.198)	(0.234)	(2,354)	(3,449)
Diversity index (D_{it})	0.323	0.145	-0.0103	-0.00739	-0.0142	-0.750	0.247	-0.328*	939.1**	-3,466
	(0.382)	(0.293)	(0.0183)	(0.0243)	(0.289)	(0.475)	(0.250)	(0.170)	(451.2)	(2,331)
Total number of sectors present in the location (M_{it})	1.815	0.890***	0.0144	0.0631**	0.109	1.599***	0.000846	0.111	50.94	-1,266
	(1.628)	(0.313)	(0.0812)	(0.0253)	(1.234)	(0.506)	(1.065)	(0.181)	(2,348)	(2,616)
Competition										
Total number of firms in the sector (N_{ijt})	0.735	-0.168	0.00431	-0.0209	0.219	-0.225	0.371	-0.0678	417.0	-4,233
	(0.499)	(0.162)	(0.0248)	(0.0132)	(0.378)	(0.262)	(0.326)	(0.0936)	(2,771)	(3,512)
Competition Index (C_{ijt})	0.276	0.273	0.00446	-0.0172	-0.106	0.327	-0.301	0.0670	-1,033	-538.2
	(0.326)	(0.205)	(0.0158)	(0.0166)	(0.247)	(0.332)	(0.214)	(0.119)	(939.0)	(1,169)

Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Instrumented agglomeration	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: All variables are in log form. Standard errors are in parenthesis. *** Indicate 1% significance level, ** indicates 5% significance level and * indicates 10% significance level. These estimation uses 5-digit industrial classification.

Table 2.7: Estimates of agglomeration and total factor productivity of firms (using M function)

	TFP (Levinhson and Petrin)			TFP (Wooldridge)		
	(1)	(2)	(3)	(4)	(5)	(6)
M function (10 kms radius)	0.0287			0.0535		
	(0.0668)			(0.138)		
M function (5 kms radius)		0.0287			0.0535	
		(0.0668)			(0.138)	
M function (2 kms radius)			0.038			0.0433
			(0.0606)			(0.124)
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Level of Geographical aggregation	10 Kms radius	5 Kms radius	2 Kms radius	10 Kms radius	5 Kms radius	2 Kms radius

Note: All variables are in log form. Standard errors are in parenthesis. *** Indicate 1% significance level, ** indicates 5% significance level and * indicates 10% significance level. These estimation uses 5- digit industrial classification.

Table 2.7a: Estimates of agglomeration and total factor productivity of firms (using M function) for

older and newer clusters

	(1)	(2)	(3)	(4)	(5)	(6)
	Older clusters	Newer clusters	Older clusters	Newer clusters	Older clusters	Newer clusters
M function (10 kms radius)	-0.18	0.11				
	(0.137)	(0.0749)				
M function (5 kms radius)			-0.18	0.11		
			(0.137)	(0.0749)		
M function (2 kms radius)					-0.0981	0.101
					(0.118)	(0.0692)
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Level of Geographical aggregation	10 Kms radius	10 Kms radius	5 Kms radius	5 Kms radius	2 Kms radius	2 Kms radius

Note: All variables are in log form. Standard errors are in parenthesis. *** Indicate 1% significance level, ** indicates 5% significance level and * indicates 10% significance level. These estimation uses 5- digit industrial classification.

Table 2.8: Decile estimates analyzing the impact of localization on total factor productivity

	TFP (Levinhson and Petrin)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	10	20	30	40	50	60	70	80	90	99
M function (5 kms radius)	-0.0466	0.00326	-0.00813	-0.0397	0.00339	0.0342	0.0651	0.0613	0.112	0.0262
	(0.0938)	(0.0267)	(0.0401)	(0.0768)	(0.0729)	(0.0793)	(0.0816)	(0.0854)	(0.0970)	(0.328)

	TFP (Wooldridge)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	10	20	30	40	50	60	70	80	90	99
M function (5 kms radius)	-0.0267	-0.00313	-0.00452	-0.0390	0.00638	0.0576	0.0637	0.0482	0.107	-0.0193
	(0.0779)	(0.0329)	(0.0425)	(0.0774)	(0.0774)	(0.101)	(0.0773)	(0.0765)	(0.0880)	(0.250)

	TFP									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	10	20	30	40	50	60	70	80	90	99
Localization										
Total size of the sector	0.00495	0.0276*	0.0195	0.0483**	0.0296	0.0339*	0.0453**	0.0176	-0.0119	-0.0595
	(0.0170)	(0.0151)	(0.0155)	(0.0192)	(0.0277)	(0.0194)	(0.0180)	(0.0182)	(0.0319)	(0.0695)
Individual Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Instrumented agglomeration	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Level of Industrial classification	5 digits	5 digits	5 digits	5 digits	5 digits	5 digits	5 digits	5 digits	5 digits	5 digits
Level of Geographical aggregation	District	District	District	District	District	District	District	District	District	District

Note: All variables are in log form. Standard errors are in parenthesis. *** Indicate 1% significance level, ** indicates 5% significance level and * indicates 10% significance level. These estimation uses 5-digit industrial classification. Deciles have been based upon TFP.

Table 2.9: Estimates of agglomeration and markups of firms (continuous space approach)

	(1)	(2)
Localization		
Total size of the sector	0.0212***	0.00870*
	(0.00534)	(0.00504)
Urbanization		
Total Employment	0.0203	0.00118
	(0.0126)	(0.00716)
Diversity Index	-0.0105	0.00821
	(0.0212)	(0.0129)
Total Number of Sectors	-0.00263	-0.00209
	(0.00204)	(0.00196)
Competition		
Total number of firms in a sector	-0.0504***	-0.0260***
	(0.00939)	(0.00985)
Competition index	-0.0029	-0.00155
	(0.00323)	(0.00325)
Industry Fixed Effects	Yes	Yes
District Fixed Effect	Yes	Yes
Level of Geographical aggregation	10 kms radius	5 kms radius

Note: All variables are in log form. Standard errors are in parenthesis. *** Indicate 1% significance level, ** indicates 5% significance level and * indicates 10% significance level. These estimation uses 5- digit industrial classification. The outcome in all specifications is markups derived using TFP estimated through Levinhson and Petrin and Wooldridge.

Table 2.10: Estimates of agglomeration and mark-ups of firms (using M function)

	Markups (Levinhson and Petrin)			Markups (Wooldridge)		
	(1)	(2)	(3)	(4)	(5)	(6)
M function (10 kms radius)	0.0182*** (0.00665)			0.00886** (0.00449)		
M function (5 kms radius)		0.0182*** (0.00665)			0.00886** (0.00449)	
M function (2 kms radius)			0.0138** (0.00590)			0.00767* (0.00399)
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Level of Geographical aggregation	10 Kms radius	5 Kms radius	2 Kms radius	10 Kms radius	5 Kms radius	2 Kms radius

Note: All variables are in log form. Standard errors are in parenthesis. *** Indicate 1% significance level, ** indicates 5% significance level and * indicates 10% significance level. These estimation uses 5- digit industrial classification. The outcome in all specifications is markups derived using TFP estimated through Levinhson and Petrin and Wooldridge.

Table 2.10a: Estimates of agglomeration and markups of firms (using M function) for older and newer

clusters

	(1)	(2)	(3)	(4)	(5)	(6)
	Older clusters	Newer clusters	Older clusters	Newer clusters	Older clusters	Newer clusters
M function (10 kms radius)	-0.00675	0.0247***				
	(0.00736)	(0.00838)				
M function (5 kms radius)			-0.00675	0.0247***		
			(0.00736)	(0.00838)		
M function (2 kms radius)					-0.00482	0.0198***
					(0.00633)	(0.00757)
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Level of Geographical aggregation	10 Kms radius	10 Kms radius	5 Kms radius	5 Kms radius	2 Kms radius	2 Kms radius

Note: All variables are in log form. Standard errors are in parenthesis. *** Indicate 1% significance level, ** indicates 5% significance level and * indicates 10% significance level. These estimation uses 5- digit industrial classification. The outcome in all specifications is markups derived using TFP estimated through Levinhson and Petrin approach

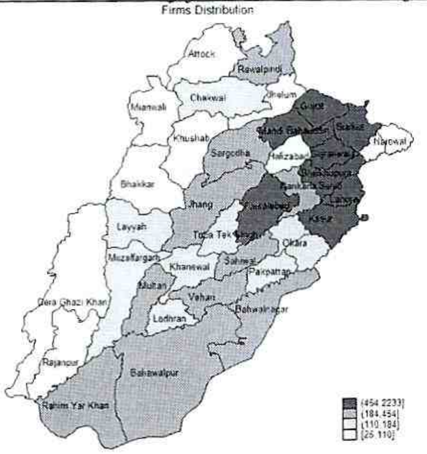
Table 2.11: Decile estimates analyzing the impact of agglomeration on mark-ups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	10	20	30	40	50	60	70	80	90	100
M	0.000438	0.000817**	0.000840*	0.000976**	0.00153	0.00133	0.00344***	0.00374**	0.00937*	0.0115
function	(0.00400)	(0.000355)	(0.000494)	(0.000426)	(0.000981)	(0.000906)	(0.00126)	(0.00186)	(0.00518)	(0.0204)

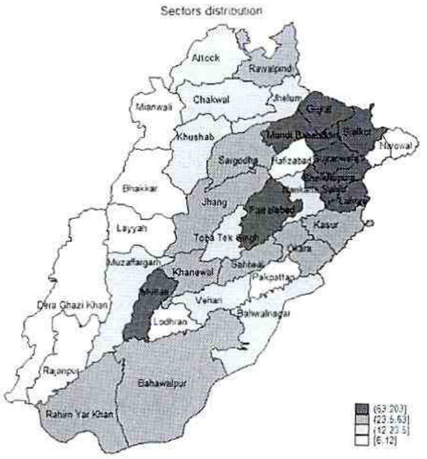
Note: Graphs have been created from decile regression estimates. Each graph shows the impact of one of the independent variables stated above on each graph on mark-ups according to distribution of mark-ups.

Figure 2.1: Distribution of firms (upper) and sectors (lower) across districts of Punjab, Pakistan.

Distribution of firms across Punjab, Pakistan



Distribution of sectors across Punjab, Pakistan



Source: Authors' calculations based on Directory of Industries 2014 for Punjab, Pakistan. The figure uses full data available and not the panel nature of the dataset used in empirical estimations.
Note: Graphs used 5-digit industrial classification to define sectors.

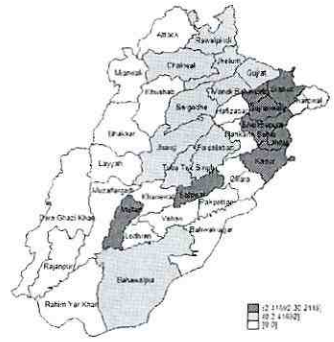
Figure 2.2: Distribution of firms from major industries across district in Punjab, Pakistan

Distribution of firms from Sports Industry

Distribution of firms from Surgical Industry

Distribution of firms from Textile Industry

Distribution of firms from Leather Industry



Distribution of firms from Electrical Equipment Industry

Distribution of firms from Food Industry

Source: Authors' calculations based on Directory of Industries 2014 for Punjab, Pakistan. The figure uses full data available and not the panel nature of the dataset used in empirical estimations.

Note: Two-digit industrial classification has been used to define Textile, Leather, Food and Electrical equipment industry and four-digit industrial classification has been used for Sports and Surgical instrument industry.

Figure 2.3: Productivity distribution of firm across older and newer cluster in 2006

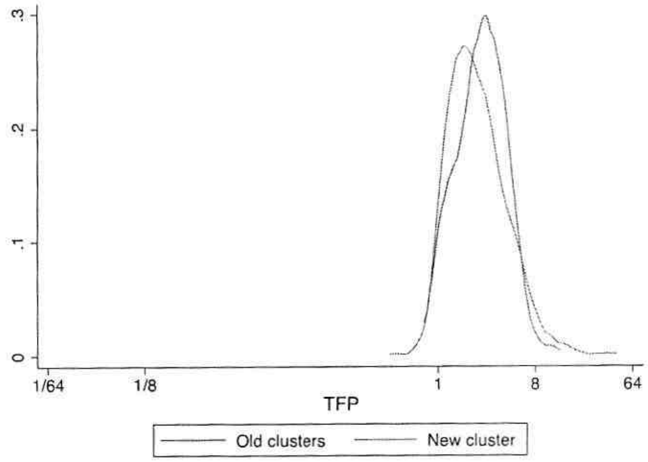
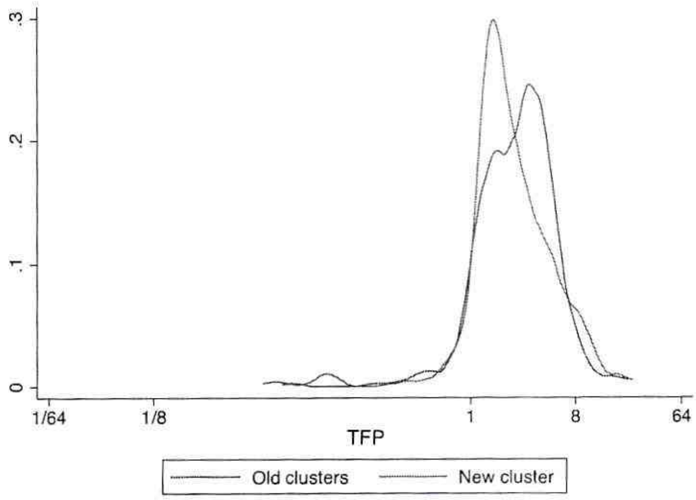


Figure 2.4: Productivity distribution of firm across older and newer cluster in 2011



References

Andersson, M., & Lööf, H. (2011). Agglomeration and productivity: Evidence from firm level data. *The Annals of Regional Science*, 46(3), 601-620.

Atkin, D., Chaudhry, A., Chaudhry, S., Khandelwal, A. K., & Verhoogen, E. (2015). Markup and cost dispersion across firms: Direct evidence from producer surveys in Pakistan. *American Economic Review Papers and Proceedings*, 105(5), 537-44.

Atkin, D., Chaudhry, A., Chaudry, S., Khandelwal, A. K., & Verhoogen, E. (2017) Organizational barriers to technology adoption: Evidence from soccer-ball producers in Pakistan. *The Quarterly Journal of Economics*, 132(3), 1101-1164.

Basant, R. (2002). Knowledge flows and industrial clusters.

Beaudry, C., & Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. Research Policy, 38 (2), 318-337.

Brooks, W. J., Kaboski, J. P., & Li, Y. A. (2016). Growth policy, agglomeration, and (the lack of competition (Working Paper No. w22947). Cambridge, MA: National Bureau ofEconomic Research.

Burki, A. A., & Khan, M. A. (2010). Spatial inequality and geographic concentration of manufacturing industries in Pakistan. The Pakistan Development Review, 49(4).

Cainelli, G., & Iacobucci, D. (2012). Agglomeration, Related Variety, and Vertical Integration. Economic Geography, 88 (3), 255-277.

Capasso, M., Stam, E., & Cefis, E. (2015). Industrial Dynamics and Economic Geography. *Regional Studies*, 49(1), 5-9.

Ciccone, A., Hall, R.E., (1996). Productivity and the Density of Economic Activity. *American Economic Review*, 86 (1), 54–70.

Ciccone, A., & Matsuyama, K. (1996). Start-up costs and pecuniary externalities as barriers to economic development. Journal of Development Economics, 49(1), 33 -59.

Combes, P. P., Magnac, T., & Robin, J. M. (2004). The Dynamics of Local Employment inFrance. *Journal of Urban Economics*, 56(2), 217-243.

Combes, P. P., Duranton, G., Gobillon, L., & Roux, S. (2010). *Estimating agglomeration economies with history, geology, and worker effects.* In Agglomeration Economics(pages 15-66). University of Chicago Press.

Combes, P. P., Duranton, G., & Gobillon, L. (2011). The identification of agglomeration economies. *Journal of Economic Geography*, 11 (2), 253-266.

De Loecker, J., & Warzynski, F. (2012). Markups and firm-level export status. *American Economic Review*, 102(6), 2437-71.

Duranton, G., & Kerr, W. R. (2015). *The logic of agglomeration* (Working Paper No. W21452). Cambridge, MA: National Bureau of Economic Research.

Duschl, M., Scholl, T., Brenner, T., Luxen, D., & Raschke, F. (2015). Industry-specific firm growth and agglomeration. *Regional Studies*, 49(11), 1822-1839.

Fafchamps, M., & Minten, B. (2002). Returns to social network capital among traders. *Oxford Economic Papers*, 54(2), 173-206.

Fafchamps, M., & El Hamine, S. (2017). Firm productivity, wages, and agglomeration externalities. *Research in Economics,* 71(2), 291-305.

Fontagné, L., & Santoni, G. (2019). Agglomeration economies and firm-level labor misallocation. Journal of Economic Geography, 19(1), 251-272.

Fujita, M., Krugman, P. R., & Venables, A. (1999). The spatial economy: Cities, regions, and international trade. MIT press.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. *Journal of Political Economy*, 100 (6), 1126-1152.

Green, E. J., & Porter, R. H. (1984). Noncooperative collusion under the imperfect price information. *Econometrica: Journal of the Econometric Society*, 87-100.

Genesove, David, and Wallace Mullin. 1998. Testing static oligopoly models: conduct and cost in the sugar industry, 1890-1914. *RAND Journal of Economics*, 29(2): 355 377.

Grillitsch, M., & Nilsson, M. (2017). Firm performance in the periphery: On the relation betweenfirm-internal knowledge and local knowledge spillovers. *Regional Studies,* 51(8), 1219- 1231.

Henderson, J. V. (2003). Marshall's scale economies. Journal of Urban Economics, 53(1), 1-28.

Hirschman, A. O. (1958), The Strategy of Economic Development, Yale University Press, New Haven.

Howell, A., He, C., Yang, R., & Fan, C. (2016). Technological relatedness and asymmetrical firm productivity gains under market reforms in China. *Cambridge Journal of Regions, Economy, and Society*, 9(3), 499-515.

Jacobs, J. (1984), *Cities and the Wealth of Nations*, Random House, New York.

Kugler, M., & Verhoogen, E. (2012). Prices, plant size, and product quality. *The Review of Economic Studies*, 79(1), 307-339.

Lang, G., E. Marcon, and F. Puech. (2015). *Distance-based measures of spatial concentration: introducing a relative density function.* Working Papers, HAL, http://EconPapers.repec.org/RePEc:hal:wpaper:hal-01082178.

Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70 (2), 317-341.

Marshall, A. (1920). *Principles of Economics: An Introductory Volume*. Royal Economic Society (Great Britain).

Marshall, Robert, and Leslie Marx. (2012). *The Economics of Collusion.* MIT Press.

Martin, P., Mayer, T., & Mayneris, F. (2011). Spatial concentration and plant-level productivity in France. *Journal of Urban Economics*, 69(2), 182-195.

Nadvi, K. (1999). Shifting ties: Social networks in the surgical instrument cluster of Sialkot, Pakistan. *Development and Change*, 30(1), 141-175.

Olley, G. S., and A. Pakes. 1996. The dynamics of productivity in the telecommunications equipment industry. *Econometrica,* 64, 1263–1297.

Overman, H. G., & Puga, D. (2010). *Labor pooling as a source of agglomeration: An empirical investigation in agglomeration economics* (Pp. 133-150). University of Chicago Press.

Rauch, J. E., & Casella, A. (2003). Overcoming informational barriers to international resource allocation: prices and ties. The Economic Journal, 113(484), 21-42.

Rigby, D. L., & Essletzbichler, J. (2002). Agglomeration economies and productivity differences in US cities. *Journal of Economic Geography*, 2(4), 407-432.

Rigby, D. L., & Brown, W. M. (2015). Who benefits from agglomeration? *Regional Studies,* 49(1), 28-43.

Rosenthal, S. S., & Strange, W. C. (2001). The determinants of agglomeration. *Journal of Urban Economics*, 50(2), 191-229.

Rosenthal, S. S., & Strange, W. C. (2003). Geography, industrial organization, and agglomeration. *Review of Economics and Statistics*, 85(2), 377-393.

Porter, M., 1990. The Competitive Advantage of Nations. Macmillan, London.

Siba, E., Söderbom, M., Bigsten, A., & Gebreeyesus, M. (2012*). Enterprise agglomeration, outputprices, and physical productivity: Firm-level evidence from Ethiopia,* (Working paper No. 2012/85). Wider Working Paper.

Widodo, W., Salim, R., & Bloch, H. (2014). Agglomeration economies and productivity growth in manufacturing industry: Empirical evidence from Indonesia. *Economic Record,* 90, 41- 58.

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservable. *Economics Letters*, 104(3), 112-114.

World Bank (2009) World Development Report 2009: *Reshaping Economic Geography.* The World Bank, Washington DC.

3. Paper II: Trade, product mix and quality of products: Evidence from the Pakistan-China Free Trade Agreement

Section 3.1: Introduction

Much of the recent trade literature has focused on the impact of changes in trade policy on firm outcomes such as product variety and quality, input use, and productivity. These changes range from unilateral trade policies, such as changes in tariffs, to broad trade agreements that involve a large number of countries, such as the North American Free Trade Agreement (NAFTA). What is generally common across the literature is a focus on one policy channel, such as the impact of import competition or export opportunity on firms. In this paper, we use a major bilateral trade agreement between China and Pakistan to test the impact of multiple channels of trade policy on firms simultaneously.

We use detailed firm-level data to (a) test the impact of this trade agreement on firm-level outcomes, and (b) disentangle the impact of greater access to export markets (due to lower foreign tariffs on exports) from the impact of greater competition in domestic markets (due to lower domestic tariffs on imported products) and greater competition in input markets (due to lower domestic tariffs on imported raw materials).

As is the case in many trade agreements between a large country and a small country, the reduction in Pakistani tariffs on Chinese goods led to a significant increase in Chinese exports to Pakistan. This, in turn, led to significantly greater competition for domestic manufacturers, which may have driven some firms out of the market and affected firms' product mix, product quality, and productivity. The evidence for other bilateral trade agreements is well documented in the literature, which finds that domestic firms may act defensively by switching product if faced with low-cost imports (Bernard, Jensen, & Schott, 2006) or may focus on their 'core competencies' by dropping their least-productive products and increasing sales per product when facing increased competition from imports (Eckel & Neary, 2010). The literature also shows that, when faced with increased competition from imports, the aggregate productivity of domestic firms rises as less productive firms exit and the remaining firms expand (see, for example, Melitz, 2003; Pavcnik, 2002).

Lower Chinese tariffs on Pakistani goods also led to higher exports from Pakistan to China, although this increase was smaller than the growth in Chinese exports to Pakistan. Higher exports to China may have also affected product mix, product quality, input use, and productivity in Pakistani firms. Firms facing greater export opportunities may have expanded their scope or changed their core products as an offensive strategy to take advantage of the export opportunities. Firms tend to move towards a few successful products and focusing on these enables them to survive and grow (Bernard, Redding, & Schott, 2011; Eckel & Neary, 2010; Iacovone & Javorcik, 2010). Changes such as these have been found to raise overall firm productivity (De Loecker, 2011).

Besides the import and export of finished goods, lower tariffs on Chinese imports gave Pakistani manufacturers increased access to cheaper imported raw materials, leading to greater competition in domestic input markets. Researchers have found a substantial increase in product variety across firms as a result of cheaper or previously unavailable imported inputs (see, for example, Amiti & Konings, 2007; Goldberg, Khandelwal, Pavcnik, & Topalova, 2010; Halpern, Koren, & Szeidl, 2015) and discovered that tariff reductions lead to the use of a greater variety of inputs as well as higher productivity (Brandt, Van Biesebroeck, Wang, & Zhang, 2017; Topalova & Khandelwal, 2011; Yu, 2015), although this impact may be stronger for importing firms as opposed to non-importing firms (Bigsten, Gebreeyesus, & Söderbom, 2016).

In addition to changes in inputs and final products, the trade literature has also analysed how firm productivity responds to changes in trade policy (Amiti & Konings, 2007; Bigsten et al., 2016; Bloom, Draca, & Van Reenen, 2016; Brandt et al., 2017; Fieler & Harrison, 2018; Pavcnik, 2002; Topalova & Khandelwal, 2011). Increased import competition can pressure domestic firms into improving their efficiency (Holmes & Schmitz, 2001; Nishimizu & Robinson, 1984) and encourage them to exploit economies of scale (Helpman & Krugman, 1985). Import competition may further lead to a reallocation of resources if high-cost producers are forced to exit the market, which frees up resources for the efficient firms that survive (Roberts & Tybout, 1991; Rodrik, 1992). The productivity gains from increased access to higher-quality intermediate inputs can also be a potential source of productivity improvement (Topalova & Khandelwal, 2011). At the same time, increasing integration with the world economy can increase firm productivity since it allows access to global finance and exposure to new goods, new methods of production, and new markets (Dornbusch, 1992; Grossman & Helpman, 1991; Romer, 1994). The evidence suggests that firms in industries exposed to greater tariff reductions increase investment in technology faster; this effect is highest in the upper-middle range of the firm size distribution (Bustos, 2011).

Much of the literature has looked at these impacts separately. We contribute to the literature in multiple ways. First, we distinguish between the different policy channels through which firm-level outcomes change, since the tariff reductions on the Pakistani and Chinese sides vary according to the industrial sector. Using detailed information on Chinese tariffs on Pakistani products as well as Pakistani tariffs on Chinese goods, we analyse the differential impact of these trade policy changes on firms in Pakistan. Since our data has details of firm-level inputs and outputs, we are able to separate the impact of lower tariffs on Chinese goods that compete with Pakistani final goods from the impact of lower tariffs on Chinese goods that serve as inputs for Pakistani firms. Pakistan is an especially interesting case since the free trade agreement led to significant decreases in Pakistani import tariffs on Chinese goods (leading to large inflows of both Chinese final goods and intermediate inputs) as well as significant decreases in Chinese import tariffs on Pakistani goods (leading to an increase in Pakistani exports to China).

So, we are able to investigate the net impact of a bilateral trade agreement on firms who were simultaneously affected by greater export opportunities, greater import competition, and cheaper imported inputs. Second, we look at the impact of a bilateral trade agreement between China and a developing country under the recently launched Belt and Road Initiative. Though the initiative includes many levels of interaction between China and developing countries, such as foreign direct investment and bilateral aid, an important element of this initiative is the promotion of trade links between China and developing countries. Up to now, the impact on firms of these enhanced trade links between China and the developing world has been understudied.

We use a newly constructed panel dataset for firms in Pakistan, drawing on data from the Census of Manufacturing Industries (CMI) from before and after the implementation of the free trade agreement (FTA). Our research adds to a thin but growing literature that disentangles the quality and productivity gains that arise from reductions on tariffs on final and intermediate inputs (Amiti & Konings, 2007; Brandt et al., 2017; Copestake, 2020; Fieler & Harrison, 2018; Lovo & Varela, 2020)—both important channels through which China's integration with the global trading system has affected firms in developing countries. Additionally, by looking at the trading relationship between China and a developing country, this paper analyses the potential benefits of cheaper and higher-quality Chinese inputs for industrial upgrading, which is broadly applicable to many developing countries.

We find that Pakistani firms added more products after the FTA as a result of lower Chinese tariffs on Pakistani goods and added fewer products because of lower Pakistani tariffs on imported Chinese inputs. We also find that firms improved their product quality because of lower Pakistan tariffs on imported Chinese inputs. Additionally, we find that firms decreased their prices and increased the quality of their inputs as a result of lower tariffs on imported Chinese inputs. Finally, we find that reductions in Chinese tariffs on Pakistani goods reduced firm-level productivity while lower tariffs on Chinese imports increased productivity, although the net effect was an overall fall in productivity.

In order to take into account heterogeneous impacts across firms, we also test the impact of different types of tariff changes on our entire sample of firms and then test to see if the impacts are different for firms that were more productive before the FTA compared to firms that were less productive. In doing so, we find that most of the impact of tariff changes on product quality, product prices, input use, and productivity is concentrated on high-productivity firms while low-productivity firms are relatively unaffected.

This article is organised as follows. Section 3.2 gives the background of the Pakistan-China FTA. In Section 3.3, we discuss the data used, and in Section 3.4, we present our methodology. Section 3.5 presents our results, and we conclude in Section 3.6.

Section 3.2: Understanding the Pakistan-China Free Trade Agreement

In November 2006, Pakistan and China signed an FTA as part of China's One Belt One Road Initiative to promote bilateral trade. As part of the agreement, China committed to eliminating tariffs on Pakistani exports to China in areas such as industrial alcohol, cotton fabrics, bed-linen, and other home textiles, marble and other tiles, leather articles, sports goods, fruits and vegetables, iron, and steel products, and engineering goods. China also committed to reducing tariffs by 50 percent on Pakistani exports of fish, dairy products, frozen orange juice, plastic products, rubber products, leather products, knitwear woven garments, and other finished textile products. At the same time, Pakistan committed to reducing tariffs on Chinese exports with a primary focus on machinery, organic and inorganic chemicals, fruits and vegetables, medicaments, and raw materials for various industrial sectors.

As part of the FTA, China committed to completely eliminating tariffs on 2,423 tariff lines, reducing tariffs by 0–5 percent on 1,338 tariff lines, granting 157 lines a 50 percent margin of preference, and granting an additional 1,768 lines a 20 percent margin of preference; China offered no concessions on 1,025 tariff lines. At the same time, Pakistan committed to completely eliminating tariffs on 2,681 tariff lines, reducing tariffs by 0–5 percent on 2,604 tariff lines, and granting a 50 percent margin of preference on 604 tariff lines and a 20 percent margin of protection on 529 tariff lines; Pakistan offered no concessions on 1,132 tariff lines[11]. The agreement was to be implemented in two five-year phases, with Phase I commencing in 2007.

Following the trade agreement, bilateral trade between Pakistan and China grew by 242 percent—nearly sixfold the growth in Pakistan's trade with the rest of the world in the same period (Afraz & Mukhtar, 2019). Figure 3.1 shows that Pakistan's exports to China rose from less than USD1 billion per year before the FTA to almost USD3 billion after the FTA, while exports from China to Pakistan grew from approximately USD4 billion a year to almost USD7 billion per year.

[Figure 3.1 near here]

The bilateral trade agreement led to tariff reductions that were started in 2007 and provides an opportunity to explore the effect of different types of tariff changes on firm-related outcomes. We explore the impact of tariff reductions through increased import competition, increased export opportunity, and greater input availability on the product scope, product quality, and productivity of Pakistani firms. The reduction in Pakistani tariffs on Chinese products increased import competition for firms in Pakistan. At the same time, the reduction in Chinese tariffs on Pakistani products created export opportunities for firms in Pakistan. Increased access to Pakistani markets not only increased import competition in Pakistan but also provided local manufacturers with both higher quality and a greater variety of inputs as well as making domestic input markets

[11] For more details on the trade agreement, please see: http://www.commerce.gov.pk/about-us/trade-agreements/pak-china-free-trade-agreement-in-goods-investment/

more competitive. Pakistan's top import categories are dominated by imported inputs for manufacturing, supporting our argument[12].

Figure 3.2 shows the relationship between the change in tariffs imposed by China on Pakistani goods and the change in Pakistani exports to China from 2006 to 2011. We observe that various sectors increased their exports to China as a result of the decline in Chinese tariffs. Some of these sectors include fish and crustaceans, salt, sulphur, earths and stone, plastering materials, ores, slag and ash, organic chemicals, plastic-related articles, leather products, and textile products.

Figure 3.3 shows the relationship between the change in the tariffs imposed by Pakistan on Chinese goods and the change in Chinese exports to Pakistan from 2006 to 2011. Here, we see that the sectors that gained the most in terms of exports to Pakistan include plastic-related articles, iron and steel, vehicles, machinery, and mechanical appliances, organic chemicals, fertilisers, and electrical machinery and equipment.

[Figure 3.2 near here]

[Figure 3.3 near here]

The figures imply that the tariff concessions made by Pakistan tended to be larger than those made by China. This, in turn, led to a greater increase not only in the number of Chinese sectors that benefitted from higher exports to Pakistan but also in total exports from China to Pakistan compared to the increase in Pakistani exports to China.

Section 3.3: Data

We use firm-level data from the CMI for before and after the implementation of

[12] Please see https://wits.worldbank.org/

the FTA between Pakistan and China (2005/06 and 2010/11)[13]. The census is conducted every five years using the sampling frames of the provincial labour departments and following a consistent methodology across years[14]. Carried out by the provincial directorates of industries, the CMI contains data on 4,000 manufacturing units, representing approximately 35–40 percent of firms in Punjab. Punjab is the largest province in Pakistan in terms of population and accounts for nearly half of all manufacturing value-added in the country. Each firm in the CMI is assigned an industry code, using the Pakistan Standard Industrial Classification of All Economic Activities (PSIC), which is in line with the International Standard Industrial Classification of All Economic Activities (ISIC). The CMI for 2011 employs the PSIC 2010 (Rev. 4), which is in line with the ISIC (Rev. 4), while the CMI for 2006 employs the PSIC 2007, which is in line with the ISIC (Rev. 3)[15].Each firm also reports a central product classification for its products, enabling us to classify each product within an industry.

The CMI collects information on production, revenues, value added, raw material costs, labour costs, energy costs, and fixed assets. It also collects information on ownership structure, firm location, district location, and industry type. The CMI also reports prices for each product produced by a firm. Using the CMI data for two periods, we created a unique panel dataset of firms that matched firms based on their unique

[13]The CMI is conducted under Sections 9 and 10 of the General Statistics Act 1975 and Sections 5 and 6 of the Industrial Statistics Act 1942 by the Pakistan Bureau of Statistics in collaboration with the provincial directorates of industries and bureaus of statistics. For more information on the Pakistan Bureau of Statistics, see http://www.pbs.gov.pk/.

[14]The CMI 2005/06 frame was enhanced using industrial directories provided by the provincial directorates of industries as well as the results of the 2001 economic census conducted by the Federal Bureau of Statistics. The annual reports of establishments listed on the stock exchanges were also used to augment the coverage

[15] For more details on PSIC classification, see:
http://www.pbs.gov.pk/sites/default/files/other/PSIC_2010.pdf.out

characteristics across years[16]. After matching the firms, we were able to create a balanced panel and an unbalanced panel of firms; for the purposes of this paper, we have used the balanced panel.

This paper also uses industry-level tariff data from the United Nations Conference on Trade and Development (UNCTAD) Trade Analysis Information System in order to assess the firm-level impact of tariff changes. The dataset has extensive information for each country, type of non-tariff measure, affected product, and partner country. We retrieved tariff-level data at the harmonised 6-digit level and then matched this data to the United Nations central product classification and ISIC Revision 3. This process enabled us to identify variations in product-level tariffs that occurred after 2006 as a result of the FTA between China and Pakistan.

Section 3.4: Methodology

We analyse the impact of the FTA by studying three potential channels through which tariff changes can affect firms: (i) Pakistani firms may be affected by increased export opportunities caused by lower Chinese tariffs on Pakistani goods, (ii) Pakistani firms may face greater competition for their final products as a result of lower Pakistani tariffs on Chinese imports, and (iii) Pakistani firms may be affected by the availability of more and cheaper inputs as well as greater competition in the input market as a result of lower Pakistani tariffs on Chinese goods that serve as potential inputs to Pakistani manufactures.

[16] We used the firm name, registration number, firm address, and telephone number to match firms across years. Most of the matching was done on the basis of names; in the case of spelling mistakes, we verified the data using other characteristics. In order to match firms as closely as possible, we initially restricted our algorithm to the data containing more than 85 percent of matches. After matching firms with the algorithm, the list was verified manually to make sure no random matches had been captured. After verifying the 85 percent of matches, we verified the remaining data manually.

The impact is measured on four sets of firm-level outcome variables: the first set of outcome variables focus on changes in product mix as measured by product add, drop and churn, which follows the work done by authors such as Boehm, Dhingra, and Morrow (in press); Goldberg et al. (2010); Grennan, Gupta, and Lederman (2018); Iacovone and Javorcik (2010); and Ma, Tang, and Zhang (2014). The second set of outcome variables focus on output quality as measured by the product quality and quality-adjusted prices, which has also been extensively studied in the literature (following Amiti & Khandelwal, 2013; Bas & Strauss-Kahn, 2015; Kugler & Verhoogen, 2012). The third set of outcome variables includes skill intensity of the workforce and indirect input linkages. This is in line with analyses by other authors such as Bastos, Silva, and Verhoogen (2018); Fieler and Harrison (2018); Javorcik (2004); and Kugler and Verhoogen (2012). Finally, we analyse whether changes in tariffs lead to an improvement in productivity (similar to Amiti & Konings, 2007; Bigsten et al., 2016; Bloom et al., 2016; Brandt et al., 2017; Fieler & Harrison, 2018; Topalova & Khandelwal, 2011).

In the next few sections, we explain how the outcome variables are measured, the construction of the variables used to measure the different tariff changes and the structure of our estimated equations.

Subsection 3.4.1: Measuring outcome variables

As discussed above, a number of authors have focused on the impact of trade liberalisation on product add, drop and churn. In order to test this in the context of the Pakistan-China FTA, we test the impact of tariff changes on the addition of new products, which is defined as the number of new products introduced at time t from time $t-1$

divided by the total number of products produced by a firm (as used by Bernard, Redding, & Schott, 2010; Boehm et al., in press; Goldberg et al., 2010; Grennan et al., 2018). We also test the impact of tariff changes on firms dropping old products, which is defined as the number of products dropped by a firm at time *t* divided by the total number of products produced at time *t*; similar measures have been used by Bernard et al. (2010); Goldberg et al. (2010); Iacovone and Javorcik (2010); and Ma et al. (2014). We then look at the impact of tariff changes on product churn, which is defined as the difference between the proportion of new products introduced by a firm and the proportion of old products no longer produced by a firm—similar in spirit to the measure of net churn used by Iacovone and Javorcik (2010) in their analysis of export varieties. Finally, we assess changes in product mix by looking at gross churn, which is defined as the sum of the proportion of new products introduced and the proportion of old products no longer produced, again in the spirit of Iacovone and Javorcik (2010).

In the third set of estimations, we analyse the impact of tariff changes on output quality and quality-adjusted prices. Here, we follow Khandelwal, Schott, and Wei (2013) to derive a measure of quality from observed quantities, prices, and the utility function. They assume that preferences are constant elasticity of substitution across horizontal varieties f and vertical quality λ_f, which enters multiplicatively with quantity q_f, such that a representative global consumer has utility:

$$U = \left(\int_{f\epsilon\Omega}\left(\lambda_f q_f\right)^{\alpha} df\right)^{1/\alpha} \tag{3.1}$$

The elasticity of substitution $\sigma \equiv 1/(1 - \alpha) > 1$ and $0 < \alpha < 1$. P is the quality-adjusted price for quality λ, elasticity of substitution σ, the demand for product f and expenditure E and can be written as:

$$q_f(\varphi) = E\lambda_f^{(\sigma-1)} P_f^{-\sigma} P^{\sigma-1} \tag{3.2}$$

Taking logs and moving prices to the left-hand side gives us:

$$lnq_f + \sigma lnp_f = (\sigma - 1)\lambda_f + \ln E + (\sigma - 1)lnP \tag{3.3}$$

Since quantity, quality and price vary over product f, firm i, industry j, and time t and expenditure and price level can vary over time, we can rewrite the equation above as:

$$lnq_{fijt} + \sigma lnp_{fijt} = (\sigma - 1)\lambda_{fijt} + \ln E + (\sigma - 1)lnP = \alpha_t + \epsilon_{fijt} \tag{3.4}$$

Adding product fixed effects to account for differing units of price or quantity across products, we get:

$$lnq_{fijt} + \sigma lnp_{fijt} = \alpha_f + \alpha_t + \epsilon_{fijt} \tag{3.5}$$

where α_f is the product fixed effect and α_t is the time fixed effect. For each firm-product-year observation, quality can be estimated as the residual from the following ordinary least squares regression with the assumption of a particular value for σ. For a given value of σ, quality is estimated as follows:

$$ln\,\hat{\lambda}_{fijt} = \hat{\epsilon}_{fijt}/(\sigma - 1) \tag{3.6}$$

Quality-adjusted prices can be calculated as follows:

$$ln\,\hat{a}_{fijt} = ln\,p_{fijt} - ln\hat{\lambda}_{fijt} \tag{3.7}$$

We follow Broda, Greenfield, and Weinstein (2006) and assume that the median elasticity of substitution for developing countries, σ, is equal to 3.3.

We also test the impact of changes in tariffs on the skill intensity of workers and higher-quality inputs. We measure skill intensity using the ratio of the wage bill of non-production workers to the wage bill of production workers, following Bernard, Redding, and Schott (2005); Bernard, Jensen, Redding, and Schott (2007); Burstein and Vogel (2010); and Tirmazee (2016). We follow the methodology used by Manova and Yu (2017) to measure the quality of inputs. They use total input expenditure on raw materials as an indicator of the quality of the firm's inputs[17]. Since we do not know whether the input is sourced domestically or imported, we use detailed input-output tables to allocate weights to expenditure on raw materials and construct a weighted average input expenditure on raw materials.

Lastly, we test the impact of tariff changes on firm productivity. We measure total factor productivity for each firm by following the methodology used by Levinsohn and Petrin (2003) who condition out serially correlated unobserved shocks to the production technology by building on Olley and Pakes (1996). Olley and Pakes (1996) use the investment decision as a proxy for unobserved productivity, while Levinsohn and Petrin (2003) use intermediate inputs as a proxy. If firms report zero investment in a significant number of cases, this casts doubt on the validity of the monotonicity condition. Hence, Levinsohn and Petrin (2003) use intermediate inputs rather than investment as a proxy

[17] We are able to exploit the rich nature of our data to obtain measures for the quality of firms' inputs used in production since our data has detailed information on the total expenditure on each input.

because firms typically report positive use of materials and energy in each year, which also implies that the monotonicity condition is more likely to hold. Provided the monotonicity condition is met and materials inputs are strictly increasing in the unobserved productivity, we can express unobserved productivity as a function of observables.

Subsection 3.4.2: Measures of trade liberalisation

As discussed above, this paper considers three possible channels to evaluate the impact of the FTA between China and Pakistan. These include: (i) greater import competition as a result of reduced Pakistani tariffs, (ii) greater export opportunities as a result of reduced Chinese tariffs, and (iii) greater competition in input markets as a result of reduced Pakistani tariffs on those Chinese goods used as inputs by firms in the Pakistani manufacturing sector. We disentangle these three channels by creating three variables that represent changes in specific tariff rates, as defined below.

In order to measure the impact of greater import competition for firms as a result of lower Pakistan tariffs, we create the variable $IMPTariff_{jt}$, which is a weighted average of Pakistan import tariffs on Chinese goods, weighted within each industry code j by Pakistan's imports from the world in the same (or nearest available) year t as the tariff.

In order to measure the impact of greater export opportunities for firms as a result of lower Chinese tariffs, we create the variable $EXTariff_{jt}$, which is a weighted average of Chinese import tariffs on Pakistani goods, weighted within each industry code j by China's imports from the world in the same (or nearest available) year t as the tariff.

Finally, in order to measure the impact of lower tariffs on inputs used by firms, we create the variable $InputTariff_{jt}$, which is the weighted sum of the tariffs on each input used by firms in industry *j*:

$$InputTariff_{jt} = \sum_k \beta jk * IMPTariff_{kt} \tag{3.8}$$

where βjk is the total value of input *k* used in industry *j* divided by the total value of all inputs used in industry *j*, calculated at the ISIC 2-digit level using input-output tables (OECD, 2004, 2009)[18]. $IMPTariff_{kt}$ is the Pakistani import tariff on Chinese goods on input *k* in time *t*.

Subsection 3.4.3: Empirical specifications

We estimate the impact of changes in tariffs on our outcome variables before and after the implementation of the FTA between Pakistan and China, by testing for the effect of tariff reductions on product addition, product drop, product churn, product quality, quality-adjusted prices, skill intensity, and productivity of firms. In our model, the outcome variables are regressed on the lagged values of our tariff-related variables, a dummy variable for periods after 2006 (after the implementation of the FTA), and an interaction between these two variables. Our model is motivated by Brandt et al. (2017) and the basic specification is:

[18] Since OECD tables are not available for Pakistan and the Asian Development Bank (ADB) does not report input-output tables for Pakistan pre-2010, we compared the ADB data available for Pakistan for 2010 with the OECD data available for the rest of the world for 2010. Since we did not find significant differences in these input-output tables for 2010, we then used the OECD input-output tables for the rest of the world for the relevant years. This data for 2004 and 2009 is available from https://stats.oecd.org/Index.aspx?DataSetCode=IOTS.

$$y_{fijt} = \alpha_1 IMPTariff_{jt-1} + \alpha_2 EXTariff_{jt-1} + \alpha_3 InputTariff_{jt-1} + \alpha_4 PostFTA_t + \alpha_5 PostFTA_t * IMPTariff_{jt-1} + \alpha_6 PostFTA_t * EXTariff_{jt-1} + \alpha_7 PostFTA_t * InputTariff_{jt-1} + \alpha_8 X_{it} + c_j + \mu_{ijt} \quad (3.9)$$

where y_{fijt} is the outcome variable (i.e., product quality or quality-adjusted price) for product f produced by firm i in sector j in time t. $IMPTariff_{jt-1}$ is the import competition variable, $EXTariff_{jt-1}$is the export opportunity variable and $InputTariff_{jt-1}$ is the foreign input availability variable. $PostFTA_t$ is a dummy variable capturing the period after the tariff reductions. In our specification, $PostFTA_t * IMPTariff_{jt-1}$ is the average effect on firms of greater import competition, $PostFTA_t * EXTariff_{jt-1}$ is the average effect on firms of greater export opportunities, and $PostFTA_t * InputTariff_{jt-1}$ is the average effect of greater access to foreign inputs. X_{it} is a vector of control variables, c_j denotes industry fixed effects and ϵ_{ijrt} is the error term. We estimate equation (3.9) for skill intensity, raw materials and total factor productivity at the firm level since we do not observe these at the product level[19]. The model is estimated using 2006 and 2011 data for the outcome variables and one-year lagged tariff data (i.e. 2005 and 2010).

To test whether tariff reductions induce the addition of products, the dropping of products, and gross and net churning between 2006 and 2011, we use the following specification:

[19]The estimations for product scope and concentration index of sales are estimated at the firm level, whereas the estimations for product quality and quality-adjusted prices are estimated at the product level.

$$y_{ijt} = \alpha_1 \Delta IMPTariff_{jt-1} + \alpha_2 \Delta EXTariff_{jt-1} + \alpha_3 \Delta InputTariff_{jt-1} + X_{it} + c_j + \mu_{ijt} \quad (3.10)$$

where y_{ijt} is an outcome variable, and $\Delta IMPTariff_{jt-1}$, $\Delta EXTariff_{jt-1}$ and $\Delta InputTariff_{jt-1}$ are the changes in import competition, export opportunity and input availability as defined above. We use tariff changes between 2005 and 2010 in our estimations.

One point to note is that the FTA resulted in a lowering of tariffs, which means that we need to interpret the coefficients in the opposite direction to their signs. Thus, if our estimated coefficient for the impact of a tariff change is negative, the direction of the FTA's impact on the outcome variable is positive (since the tariffs were decreased). So, positive coefficients point to decreases in the outcome variable as a result of lower tariffs and negative coefficients point to increases in the outcome variable as a result of lower tariffs.

In order to test for heterogeneous impacts of tariff reductions on firm-related outcomes, we divide our sample of firms into high-productivity and low-productivity firms and test the specifications above for each subsample. Firms are classified into two groups based on their productivity pre-FTA.

Subsection 3.4.4: Products, prices and productivity before and after the FTA

Before moving to our empirical results, it is useful to look at our outcome variables before and after the implementation of the FTA.

Table 3.1 shows how firms altered their product mix after the FTA was implemented. We see that 29.66 percent of firms reduced their product scope from 2006 to 2011 while 15.28 percent increased their product scope from 2006 to 2011; for the remaining 55.06 percent of firms, product scope remained unchanged. Table 3.2 looks at firm behaviour at the product level. We find that 55 percent of firms kept their product scope unchanged; 47 percent added at least one product to their product scope and 54 percent dropped at least one product from their existing product mix.

[Table 3.1 near here]

[Table 3.2 near here]

Figure 3.4 looks at what happened to product quality before and after the FTA was implemented for all firms and also for high-productivity and low-productivity firms. On the whole, product quality increased after the FTA, though this was driven by a significant increase in product quality for high-productivity firms while product quality fell for low-productivity firms. Figure 3.5 looks at product prices before and after the implementation of the FTA and we see that product prices increased for both low-productivity and high-productivity firms after the trade agreement.

[Figure 3.4 near here]

[Figure 3.5 near here]

Finally, Figure 3.6 shows the total factor productivity distribution of firms before and after the implementation of the FTA. Here, we see that mean firm-level productivity decreased slightly after the implementation of the FTA, which is driven by a fall in the mean productivity of high-productivity firms, although the mean productivity of low-productivity firms shows a slight increase.

[Figure 3.6 near here]

Section 3.5: Results

We begin by testing the impact of tariff reductions on the proportion of products added and products dropped, and on whether there is a churn in firm-level products. Table 3.3 presents our results for the impact of export opportunity, import competition, and increased access to more competitive input markets on the addition of products (column 1), the dropping of products (column 2), and product churn (columns 3 and 4).

The results in column 1 show that firms added more products as a result of lower Chinese tariffs on Pakistani goods and added fewer products because of lower Pakistani tariffs on imported Chinese inputs. In column 2, our results show that reductions in tariffs on imports had a significant impact on firms' decision to drop a product, that is, there was a significant reduction in the number of products dropped by firms in response to higher import competition that resulted from lower tariffs on imported goods. We did not find any significant impact of greater export opportunities (as a result of lower Chinese tariffs on Pakistani exports) and more competitive input markets (as a result of lower tariffs on imported Chinese inputs) on firms' decision to drop products after the implementation of the FTA.

Table 3.3 also provides findings from estimations analysing the impact of export opportunity, import competition, and increased access to imported inputs on gross product churn (column 3) and net product churn (column 4). We find that import competition has a negative effect on gross product churning or that gross product churn is less likely among firms as a result of greater import competition (which is in line with the impact of import competition on the dropping of products). We do not find any significant evidence of the impact of export opportunity and more competitive input markets on firms' gross churning of products. In column 4, we find that the higher import competition post-FTA increases the likelihood of net churn (which is, again, in line with the result that firms drop fewer products as a result of greater import competition). We also find that increased access to foreign inputs has a marginally significant impact on net churning.

Our results are consistent with recent models of multi-product firms, which find that firms adjust their product scope in response to changes in trade costs (see Iacovone & Javorcik, 2010). However, our results differ in that the literature identifies the impact of greater export opportunities on product scope while we find that, in the case of Pakistan's trade with China, product scope is affected by greater import competition.

We then present disaggregated results for high-productivity firms (columns 5 to 8) and low-productivity firms (columns 9 to 12). We find that high-productivity firms add more products and are more likely to churn their products as a result of greater export opportunities. Our results also suggest that low-productivity firms add fewer products and churn less because of lower tariffs on imported inputs.

[Table 3.3 near here]

Table 3.4 provides findings on the impact of export opportunity, import competition, and more competitive input markets on firm-level product quality (column 1) and quality-adjusted prices (column 2). For the full sample of firms, we do not find any impact of increased export opportunities and import competition on product quality post-FTA (column 1). This suggests that reductions in Chinese tariffs on Pakistani goods did not induce firms to alter their product quality. On the other hand, we do find that firms improved their product quality because of lower Pakistani tariffs on imported Chinese inputs; in other words, reductions in tariffs on imported inputs and more competitive input markets were accompanied by greater product quality. Our results are in line with the literature that finds that tariff reductions or trade liberalisation policies induce changes at the firm level through changes in input markets. This literature is driven by empirical evidence that has found that firms may take advantage of input trade liberalisation to upgrade the quality of inputs in order to upgrade the quality of their products (Bas & Strauss-Kahn, 2015; Kugler & Verhoogen, 2012). Our results are also in line with recent research on firms in India that finds that higher import competition in input markets gives firms access to cheaper and higher-quality inputs, which improves product quality (Copestake, 2020).

Column 2 presents our results for the impact of the FTA on quality-adjusted prices. We do not find any significant impact of increased export opportunities and greater import competition on quality-adjusted prices. We do find that more competitive input markets have a significant impact on quality-adjusted prices, that is, quality-adjusted prices decreased significantly for firms that used inputs on which import tariffs were lower. These results are in line with the recent literature, which demonstrates that lower input tariffs reduce output prices (see De Loecker, Goldberg, Khandelwal, & Pavcnik, 2016). The recent evidence from India suggests that higher import competition in input markets increases access to cheaper and higher-quality inputs, which lowers quality-adjusted prices (Copestake, 2020).

Again, we perform the same analysis for both high-productivity and low-productivity firms (columns 3 and 4 for high-productivity firms and columns 5 and 6 for low-productivity firms). We find that high-productivity firms improved the quality of their products and decreased quality-adjusted prices due to competition in input markets.

This suggests that our aggregate results are driven mainly by high-productivity firms. We find that low-productivity firms did not improve their quality of products or change quality-adjusted prices due to changes in input tariffs. However, we find that low-productivity firms lowered their prices as a result of greater export opportunities.

[Table 3.4 near here]

Table 3.5 presents our results for the relationship between tariff changes and changes in skill intensity and input market costs. In particular, we test the impact of tariff changes on skill intensity (column 1) and raw material costs (column 2). We do not find any significant impact of increased export opportunities and greater import competition on skill intensity and raw material costs. However, we do find that raw material costs increased as a result of lower tariffs on imported inputs, implying that firms improved the quality of their raw materials as a result of more competitive input markets. This may be because firms that were attempting to increase the quality of their output switched to higher-quality (and more expensive) raw materials.

Again, we divide our sample into high-productivity and low-productivity firms (columns 3 and 4 for high-productivity firms and columns 5 and 6 for low-productivity firms). We find that the aggregate results of the impact of tariff changes on skill intensity input and raw material costs are driven by high-productivity firms: high-productivity firms increased their skill intensity and raw material quality as a result of lower tariffs on their inputs. Additionally, we find that the skill intensity in high-productivity firms increased as a result of greater export opportunities.

These results are consistent with the literature on developing countries, which documents large increases in the demand for skilled workers, following trade liberalisation (Goldberg & Pavcnik, 2004, 2007). Our results are also consistent with the well-established literature on quality and studies utilizing O-ring assumptions that the production of higher-quality output requires more skilled labour (see Fieler & Harrison, 2018; Khandelwal, 2010; Kremer, 1993; Schott, 2004; Verhoogen, 2008) and higher-quality inputs (see Bastos et al., 2018; Kugler & Verhoogen, 2012; Manova & Zhang, 2012).

[Table 3.5 near here]

Finally, Table 3.6 presents the results for the relationship between total factor productivity and changes in tariffs. We find evidence of the impact of increased export opportunities and greater import competition on total factor productivity (column 1)[20]. Reductions in Chinese tariffs on Pakistani exports reduced firms' productivity, but lower Pakistani tariffs on competing Chinese imports induced firms to improve their productivity after the implementation of the FTA. Our results further suggest that competition in the input market had no significant effect on total factor productivity. This is in contrast to the literature, which finds that export opportunities enable firms to explore new opportunities and become more productive (Brandt et al., 2017; Fieler & Harrison, 2018) or that reduced tariffs on imported inputs induce productivity improvements (Bigsten et al., 2016).

Our findings for high-productivity firms (column 2) and low-productivity firms (column 3) show that greater import competition (lower Pakistani tariffs on Chinese imports) increased productivity in high-productivity firms while greater access to Chinese markets (lower Chinese tariffs on Pakistani exports) decreased productivity, implying that our aggregate results are driven mostly by high-productivity This goes against the

[20] We also control for product churn in the productivity estimations to see if our result changes.

usual expectation that lower foreign tariffs on exports tend to increase productivity and one possible explanation could that there may be greater product churn in high productivity firms leading to temporary decreases in TFP as the adjustments are made. None of the tariff channels had an impact on the total factor productivity of low-productivity firms.

[Table 3.6 near here]

We carry out a final exercise to estimate the net impact on the productivity of the different tariff changes. We do so because our results above show that the impacts of different tariff changes may move in opposite directions and the actual tariff changes that occurred may differ according to the type of tariff. Our estimated net impact of the different types of tariff changes on the productivity of Pakistani firms is given in Table 3.7. Here, we see that, while reductions in import tariffs had a positive impact on productivity, the net impact of tariff changes on productivity was negative because of the negative impact of reduced Chinese tariffs on Pakistani exports. Again, when we break this analysis down by firms' initial productivity level, we find that the net negative impact on productivity is significant for high-productivity firms while the net impact for low-productivity firms is insignificant.

[Table 3.7 near here]

Subsection 3.5.1: Discussion of results

Our paper provides key insights into the impact on firms of the FTA between China and Pakistan. First, we find that lower tariffs induce firms to alter their product scope by adding more products and dropping fewer products; for high-productivity firms, this is induced by lower Chinese tariffs on Pakistani exports and lower tariffs on Chinese imports to Pakistan, while for low-productivity firms, it is induced by lower tariffs on imported Chinese inputs.

Second, we find that reduced tariffs on imported Chinese inputs induce firms to improve the quality of their products and reduce their prices; this result is driven by high-productivity firms, which improve quality and reduce their prices as a result of lower tariffs on Chinese inputs, while the prices and quality of output for low-productivity firms are unaffected by changes in tariffs. In terms of skill intensity and raw material costs, lower tariffs on imported Chinese inputs, again, affect only high-productivity firms; skill intensity and raw material costs go up for high-productivity firms as a result of lower tariffs on Chinese inputs while low-productivity firms are unaffected by changes in tariffs.

Finally, we find that the productivity of higher-productivity firms is positively affected by lower import tariffs on Chinese goods but negatively affected by lower Chinese tariffs on Pakistani exports. Again, the productivity of low-productivity firms is unaffected by changes in tariffs that occurred as a result of the FTA.

The results point to a situation where high-productivity firms take advantage of increased access to Chinese markets by focusing on their existing product mix or adjusting their product mix. These firms also improve the quality of their output and reduce their prices as they try and take advantage of this greater access to Chinese markets. As they do so, they upgrade their labour force and use better raw materials. However, as they move into producing higher-quality goods to take advantage of increased access to Chinese markets, their productivity decreases. On the other hand, low-productivity firms start by focusing on their core products as a result of lower tariffs but are generally unaffected by the tariff changes that occurred as a result of the FTA.

Section 3.6: Conclusion

While our understanding of how trade policies affect firms is evolving rapidly, researchers have tended to look at a limited number of channels through which changes in trade policy can affect firm-level decisions. We have used the FTA between Pakistan and China to test simultaneously how changes in trade policies have affected Pakistani firms. In particular, we have tested to see how lower Chinese tariffs on Pakistani goods and lower Pakistani tariffs on Chinese goods affected the products and productivity of firms.

We found that the opening of Chinese markets to Pakistani firms led mostly to changes in those firms that were more productive before the implementation of the FTA. The high-productivity firms improved the quality of their products and charges lower prices but experienced a fall in productivity. On the other hand, firms that started out as less productive before the FTA tended to be unaffected by the various tariff changes.

In addition to illustrating the net impact of trade liberalisation in a developing country context, our paper also shows how a developing country is affected by a trade agreement with China. As China continues to cultivate economic ties with developing countries, it is important to gauge the economic impact of wide-ranging trade agreements on firms in these countries.

Table 3.1. The direction of change in product scope from 2006 to 2011

Direction of change in Product scope	Number of firms	Percentage of firms
Reduced Scope	299	29.66
Unchanged Scope	555	55.06
Increased Scope	154	15.28

Note: The direction of change in product scope from 2006 to 2011 is defined as the change in the number of products produced from 2006 to 2011 classified according to three categories: 'reduced scope' is defined as whether a firm's product scope was reduced from 2006 to 2011, 'unchanged scope' is defined as whether a firm produced the same number of products from 2006 to 2011, and 'increased scope' is defined as whether a firm's product scope increased from 2006 to 2011.

Source: Authors' calculations based on data from the Census of Manufacturing Industries for Punjab, 2011 and 2006.

Table 3.2. Product add, drop, churn from 2006 to 2011

	Number of firms	Percentage of firms
Add	481	47.72
Drop	551	54.66
Churn	374	37.1

Note: 'Add' is a dummy variable equal to 1 if and only if firm i does not produce a product at time t but does at time $t + 1$. 'Drop' is a dummy variable equal to 1 if and only if firm i produces a product at time t but does not at time $t + 1$. 'Churn' is a dummy variable equal to 1 if two conditions are satisfied for firm i: (i) that firm i does not produce a product at time t but does at time $t + 1$, and (ii) that firm i produces a product at time t but does not at time $t + 1$. We have information for 1,008 firms for scope in both years.

Source: Author's calculations, based on data from the Census of Manufacturing Industries for Punjab, 2011 and 2006.

Table 3.3. Impact of import competition, export opportunity, and competition in input markets on product churn

a. Full sample

	Add	Drop	Gross churn	Net churn
	(1)	(2)	(3)	(4)
ΔIMPTariff	0.00017	0.0257***	0.0258**	-0.0255***
	(0.0031)	(0.0096)	(0.0110)	(0.0091)
ΔEXTariff	-0.0039***	0.0046	0.0006	-0.0086
	(0.0015)	(0.0054)	(0.0058)	(0.0053)
ΔInputTariff	0.0489***	-0.0362	0.0127	0.0851*
	(0.0148)	(0.0541)	(0.0610)	(0.0507)
N	1,008	1,008	1,008	1,008
R-sq.	0.193	0.075	0.093	0.068
Controls	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes

b. High-productivity firms

	Add	Drop	Gross churn	Net churn
	(1)	(2)	(3)	(4)
ΔIMPTariff	0.0022	0.0251*	0.0273	-0.0229*
	(0.0042)	(0.0146)	(0.0172)	(0.0131)
ΔEXTariff	-0.0056***	0.0079	0.0022	-0.0136**
	(0.0016)	(0.0061)	(0.0065)	(0.0061)
ΔInputTariff	0.0312	-0.0696	-0.0384	0.101
	(0.0194)	(0.0815)	(0.0889)	(0.0783)
N	646	646	646	646
R-sq.	0.214	0.086	0.097	0.086
Controls	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes

c. Low-productivity firms

	Add	Drop	Gross churn	Net churn
	(1)	(2)	(3)	(4)
ΔIMPTariff	-0.0051	0.0185	0.0134	-0.0236*
	(0.0052)	(0.0134)	(0.0158)	(0.0128)
ΔEXTariff	0.0010	-0.0071	-0.0060	0.0081
	(0.0038)	(0.0124)	(0.0143)	(0.0116)
ΔInputTariff	0.0813***	-0.0088	0.0725	0.0901*
	(0.0247)	(0.0633)	(0.0792)	(0.0545)
N	362	362	362	362
R-sq.	0.214	0.098	0.129	0.086
Controls	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses. Fixed effects (FE) include industry and year fixed effects. 'Net churn' is the difference between new products produced and old products for which production has ceased (Iacovone & Javorcik, 2010). New products are defined as the number of new products introduced at time t from time $t-1$ divided by the total number of products produced by a firm. We define product destruction as the number of products that cease to be produced at time t divided by the total number of products produced at time t. 'Gross churn' is the sum of new products introduced and old products for which production has ceased (Iacovone & Javorcik, 2010). $IMTariff_{jt-1}$ is the import competition effect measured by the change in the lagged tariff. $EXTariff_{jt-1}$ is the export opportunity measured by the change in the lagged tariff and $InputTariff_{jt-1}$ is the import competition in the input market measured by the change in the lagged tariff. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Source: Authors' calculations based on data from the Census of Manufacturing Industries for Punjab, 2011 and 2006.

Table 3.4. Impact of import competition, export opportunity and competition in input markets on product quality and quality-adjusted prices

	Full sample		High-productivity firms		Low-productivity firms	
	Quality	Quality-adjusted prices	Quality	Quality-adjusted prices	Quality	Quality-adjusted prices
	(1)	(2)	(3)	(4)	(5)	(6)
IMPTariff	-0.00253	0.000808	0.0221	-0.00288	-0.00602	0.000165
	(0.0159)	(0.00669)	(0.0194)	(0.00771)	(0.0297)	(0.0142)
EXTariff	-0.0379	0.0121	-0.0836**	0.0264***	0.0195	-0.00930
	(0.0253)	(0.00803)	(0.0325)	(0.0101)	(0.0352)	(0.0156)
InputTariff	0.157*	-0.0549*	-0.00511	-0.0629*	0.148	-0.0397
	(0.0847)	(0.0297)	(0.105)	(0.0347)	(0.156)	(0.0496)
PostFTA	1.362**	0.643**	2.374***	0.535*	-0.263	0.644
	(0.684)	(0.265)	(0.794)	(0.305)	(1.477)	(0.446)
IMPTariff*PostFTA	0.00158	0.00489	-0.00969	0.00847	-0.00583	0.00163
	(0.0138)	(0.00469)	(0.0168)	(0.00574)	(0.0208)	(0.00867)
EXTariff* PostFTA	0.0136	-0.00191	0.0472*	-0.0134*	-0.0285	0.0218**
	(0.0202)	(0.00626)	(0.0256)	(0.00778)	(0.0321)	(0.0102)
InputTariff*PostFTA	-0.247**	0.0774**	-0.284**	0.0966**	-0.0650	0.0547
	(0.0985)	(0.0372)	(0.111)	(0.0422)	(0.224)	(0.0636)
N	3,226	3,226	2,473	2,473	753	753
R-sq.	0.638	0.759	0.623	0.709	0.728	0.828
Controls	Yes	Yes	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses. Fixed effects (FE) include firm and year fixed effects. Firm product quality and quality-adjusted prices are derived following Khandelwal, Schott, and Wei (2013). is the average effect of import competition in the longer run, i.e., after five years of tariff reduction. is the average effect of export opportunity in the long run and is the average effect of import competition in the input market. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Source: Authors' calculations based on data from the Census of Manufacturing Industries for Punjab, 2011 and 2006.

Table 3.5. Impact of import competition, export opportunity, and competition in input markets on skill intensity and raw material costs

	Full sample		High-productivity firms		Low-productivity firms	
	Skill intensity	Total raw material costs	Skill intensity	Total raw material costs	Skill intensity	Total raw material costs
	(1)	(2)	(3)	(4)	(5)	(6)
IMPTariff	-0.0812*	-1,728*	-0.106***	-2,795**	-0.204	941.0
	(0.0417)	(938.1)	(0.0372)	(1,146)	(0.192)	(1,278)
EXTariff	-0.0592	-2,525	0.0126	-2,531*	-0.154	-2,856
	(0.100)	(1,702)	(0.0777)	(1,523)	(0.170)	(3,363)
InputTariff	0.322	24,468**	1.253**	29,745***	-2.369	9,323
	(1.035)	(9,988)	(0.633)	(10,378)	(2.434)	(17,473)
PostFTA	5.033	210,610***	26.11***	192,253**	-39.38	152,157*
	(14.75)	(70,265)	(9.168)	(77,039)	(32.88)	(91,390)
IMPTariff*PostFTA	0.412	85.40	-0.0432	1,937	1.178	-1,669
	(0.417)	(1,548)	(0.0759)	(2,249)	(0.884)	(2,041)
EXTariff* PostFTA	-0.0839	1,125	-0.175*	1,177	0.0549	1,198
	(0.130)	(1,797)	(0.0965)	(1,584)	(0.246)	(3,619)
InputTariff*PostFTA	-0.613	-28,431***	-2.461**	-28,617***	3.815	-17,064
	(1.443)	(9,456)	(1.076)	(9,973)	(3.658)	(16,459)
N	1,604	2,099	963	1,269	641	830
R-sq.	0.128	0.079	0.188	0.068	0.262	0.198
Controls	Yes	Yes	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses. Fixed effects (FE) include firm and year fixed effects. The skill intensity ratio is measured by the ratio of the wage bill of non-production to the wage bill of production workers, following Bernard, Redding, and Schott (2005). Raw material cost is the sum of raw material costs that a firm incurred in a given year. is the average effect of import competition in the longer run, i.e., after five years of tariff reduction. is the average effect of export opportunity in the long run and is the average effect of import competition in the input market. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Source: Authors' calculations based on data from the Census of Manufacturing Industries for Punjab, 2011 and 2006.

Table 3.6. Impact of import competition, export opportunity and competition in input markets on total factor productivity

	Full sample	High-productivity firms	Low-productivity firms
	Total factor productivity		
	(1)	(2)	(3)
IMPTariff	-0.00727	-0.00312	-0.00832
	(0.00555)	(0.00704)	(0.00959)
EXTariff	-0.0253***	-0.0633***	0.0136
	(0.00914)	(0.0119)	(0.0136)
InputTariff	-0.0417	0.0201	-0.0983
	(0.0583)	(0.0656)	(0.115)
PostFTA	-0.428	0.0180	-0.866
	(0.536)	(0.578)	(1.101)
IMPTariff*PostFTA	-0.0151*	-0.0380***	0.0127
	(0.00797)	(0.0103)	(0.0124)
EXTariff* PostFTA	0.0309***	0.0636***	-0.00321
	(0.0104)	(0.0127)	(0.0180)
InputTariff*PostFTA	0.0485	-0.0158	0.0949
	(0.0776)	(0.0830)	(0.158)
N	2,266	1,312	954
R-sq.	0.253	0.247	0.2
Controls	Yes	Yes	Yes
FE	Yes	Yes	Yes

Note: Standard errors in parentheses. Fixed effects (FE) include firm and year fixed effects. Total factor productivity is measured using the semi-parametric production function approach following Levinsohn and Petrin (2003). is the average effect of import competition in the longer run, i.e., after five years of tariff reduction. is the average effect of export opportunity in the long run and is the average effect of import competition in the input market. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

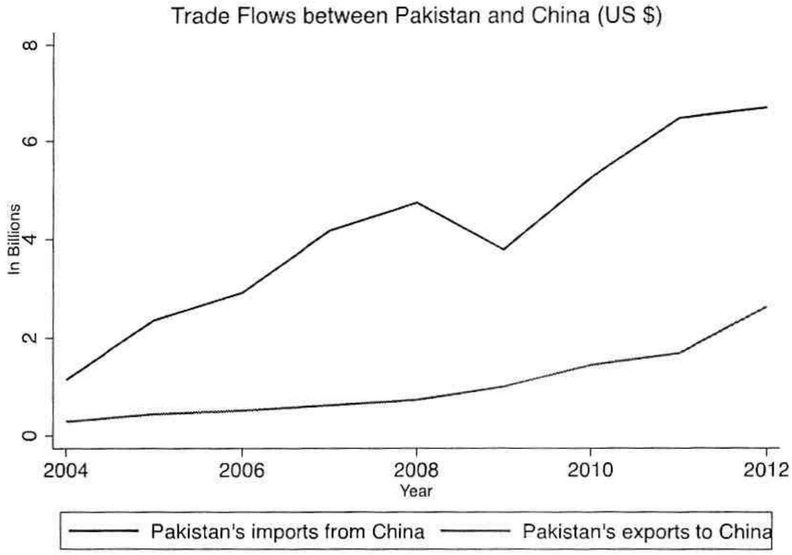
Source: Authors' calculations based on data from the Census of Manufacturing Industries for Punjab, 2011 and 2006.

Table 3.7. The net impact on total factor productivity due to changes in tariff

	Full Sample	High productivity firms	Low productivity firms
	(1)	(2)	(3)
IMPTariff	-0.031	-0.112	0
EXTariff	0.117	0.388	0
InputTariff	0	0	0
Net Impact	0.085	0.277	0

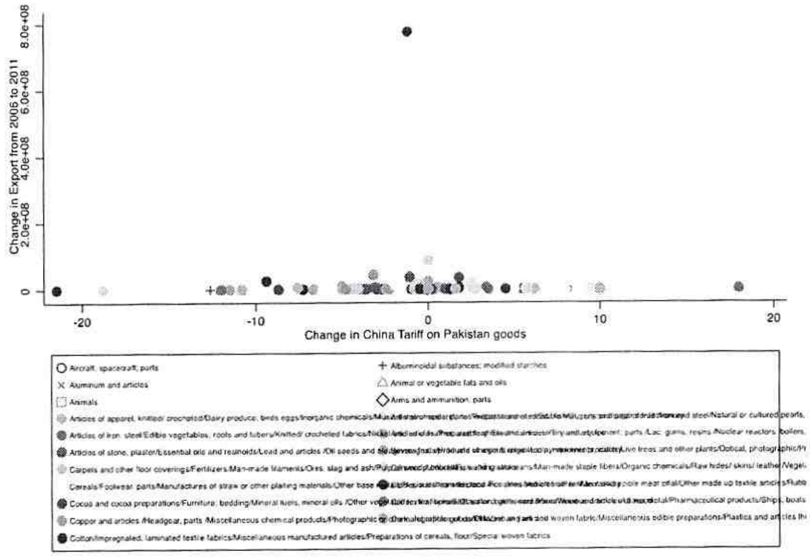
Source: Authors' calculations based on results in table 3.6

Figure 3.1: Trade flows between Pakistan and China from 2004 to 2012



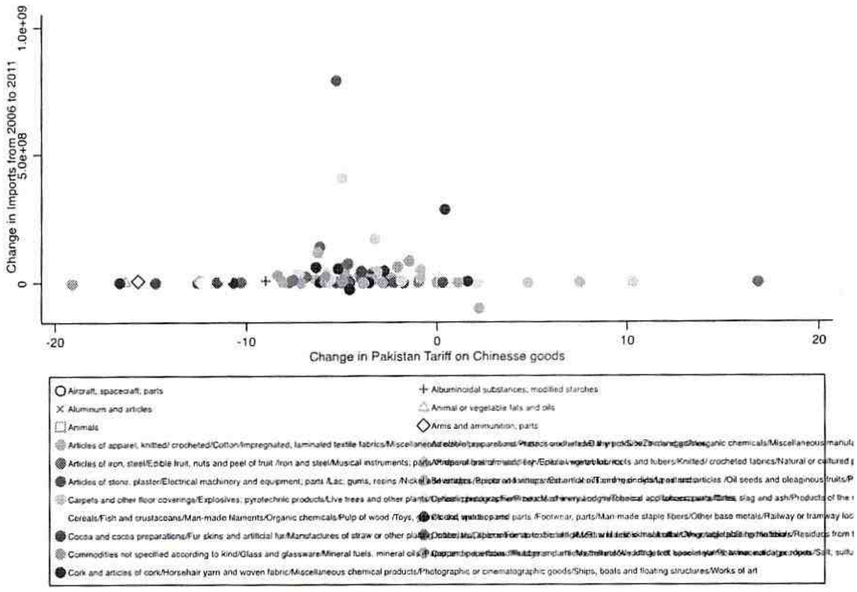
Source: Author's calculations based upon UN Comtrade data from 2004 to 2012 for Pakistan

Figure 3.2: Change in Pakistani exports to China from 2006 to 2011



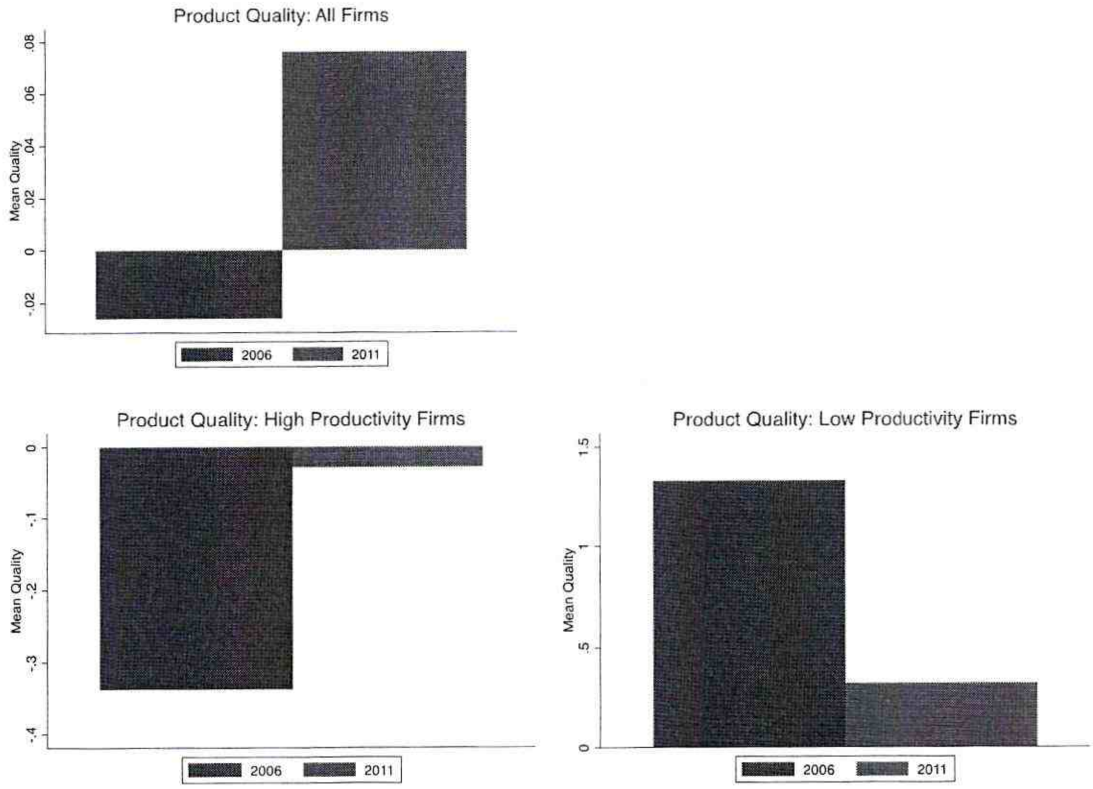
Source: Author's calculations based upon UN Comtrade and UNCTAD Trade Analysis Information System (TRAINS) data from 2006 to 2011 for Pakistan

Figure 3.3: Change in Chinese exports to Pakistan from 2006 to 2011



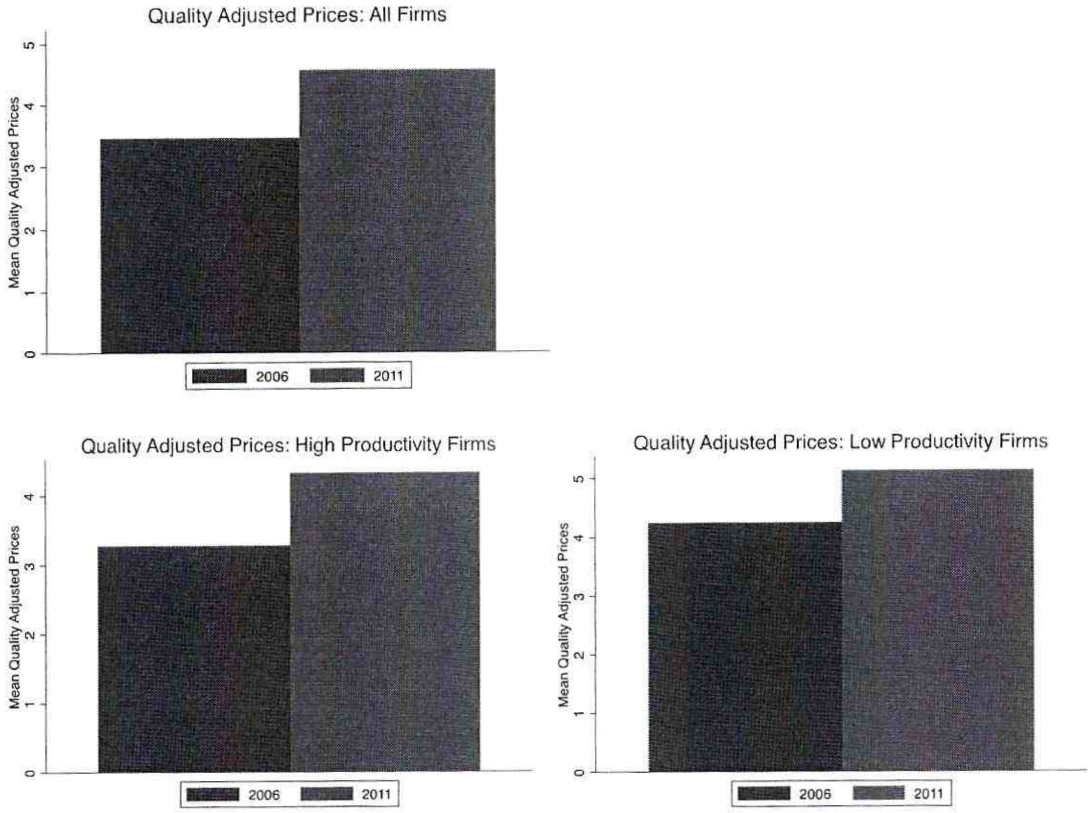
Source: Author's calculations based upon UN Comtrade and UNCTAD Trade Analysis Information System (TRAINS) data from 2006 to 2011 for Pakistan

Figure 3.4: Quality of products produced by firms



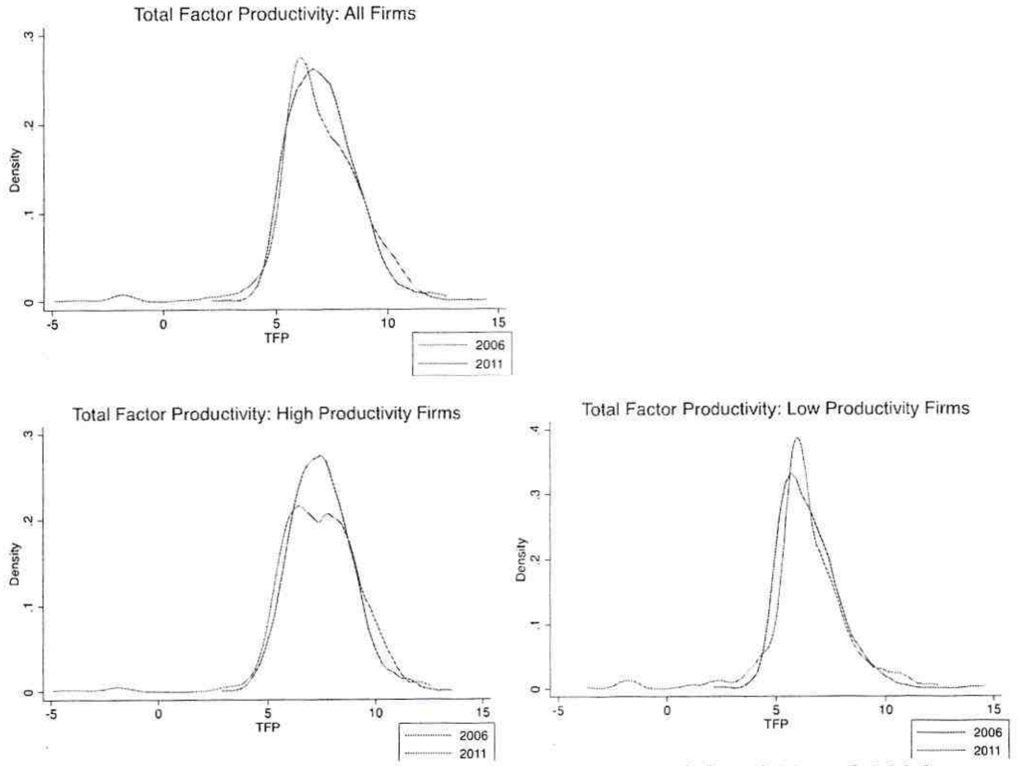
Source: Authors' own calculations based on CMI Punjab, 2011 and 2006.
Note: Firms' product quality is derived following the methodology of Khandelwal, Schott and Wei (2013).

Figure 3.5: Quality-adjusted prices of products produced by firms



Source: Authors' own calculations based on CMI Punjab, 2011 and 2006.
Note: Firm's product quality-adjusted prices are derived following the methodology of Khandelwal, Schott and Wei (2013).

Figure 3.6: Productivity distributions of firms in 2006 and 2011



Source: Authors' own calculations based on CMI Punjab, 2011 and 2006.
Note: Total factor productivity is measured using a semi-parametric production function approach following the methodology of Levinsohn and Petrin (2003).

References

Afraz, N., & Mukhtar, N. (2019). China Pakistan free trade agreement Phase 2: A preliminary analysis (Report No. R1903). Karachi: Pakistan Business Council and Consortium for Development Policy Research.

Amiti, M., & Khandelwal, A. K. (2013). Import competition and quality upgrading. *The Review of Economics and Statistics*, *95*(2), 476–490.

Amiti, M., & Konings, J. (2007). Trade liberalization, intermediate inputs, and productivity: Evidence from Indonesia. *American Economic Review*, *97*(5), 1611–1638.

Bas, M., & Strauss-Kahn, V. (2015). Input-trade liberalization, export prices, and quality-upgrading. *Journal of International Economics*, *95*(2), 250–262.

Bastos, P., Silva, J., & Verhoogen, E. (2018). Export destinations and input prices. *American Economic Review*, *108*(2), 353–392.

Bernard, A. B., Jensen, J. B., & Schott, P. K. (2006). Survival of the best fit: Exposure to low wage countries and the (uneven) growth of US manufacturing plants. *Journal of International Economics*, *68*(1), 219–237.

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2007). Firms in international trade. *Journal of Economic Perspectives*, *21*(3), 105–130.

Bernard, A. B., Redding, S., & Schott, P. K. (2005). *Factor price equality and the economies of the United States* (Discussion Paper No. 696). London: Centre for Economic Performance.

Bernard, A. B., Redding, S. J., & Schott, P. K. (2010). Multiple-product firms and product switching. *American Economic Review*, *100*(1), 70–97.

Bernard, A. B., Redding, S. J., & Schott, P. K. (2011). Multiproduct firms and trade liberalization. *The Quarterly Journal of Economics, 126*(3), 1271–1318.

Bigsten, A., Gebreeyesus, M., & Söderbom, M. (2016). Tariffs and firm performance in Ethiopia. *Journal of Development Studies, 52*(7), 986–1001.

Bloom, N., Draca, M., & Van Reenen, J. (2016). Trade-induced technical change? The impact of Chinese imports on innovation, IT, and productivity. *The Review of Economic Studies, 83*(1), 87–117.

Boehm, J., Dhingra, S., & Morrow, J. (in press). The comparative advantage of firms. *Journal of Political Economy.*

Brandt, L., Van Biesebroeck, J., Wang, L., & Zhang, Y. (2017). WTO accession and performance of Chinese manufacturing firms. *American Economic Review, 107* (9), 2784–2820.

Broda, C., Greenfield, J., & Weinstein, D. (2006). *From groundnuts to globalization: A structural estimate of trade and growth* (Working Paper No. 12512). Cambridge, MA: National Bureau of Economic Research.

Burstein, A., & Vogel, J. (2010). *Globalization, technology, and the skill premium: A quantitative analysis* (Working Paper No. 16459). Cambridge, MA: National Bureau of Economic Research.

Bustos, P. (2011). Trade liberalization, exports, and technology upgrading: Evidence on the impact of MERCOSUR on Argentinian firms. *American Economic Review, 101*(1), 304–340.

Copestake, A. (2020). *Multi-product firms, networks, and quality-upgrading: Evidence from China in India* (Working Paper No. 1659). San Rafael, CA: Forum for Research in Empirical International Trade.

De Loecker, J. (2011). Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. *Econometrica*, *79*(5), 1407 1451.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., & Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica*, *84*(2), 445–510.

Dornbusch, R. (1992). The case for trade liberalization in developing countries. *Journal of Economic Perspectives*, *6*(1), 69–85.

Eckel, C., & Neary, J. P. (2010). Multi-product firms and flexible manufacturing in the global economy. *The Review of Economic Studies*, *77*(1), 188–217.

Fieler, A. C., & Harrison, A. (2018). *Escaping import competition and downstream tariffs* (Working Paper No. 24527). Cambridge, MA: National Bureau of Economic Research.

Goldberg, P. K., Khandelwal, A. K., Pavcnik, N., & Topalova, P. (2010). Multiproduct firms and product turnover in the developing world: Evidence from India. *The Review of Economics and Statistics*, *92*(4), 1042–1049.

Goldberg, P. K., & Pavcnik, N. (2004). Trade, inequality, and poverty: What do we know? Evidence from recent trade liberalization episodes in developing countries. *Brookings Trade Forum*, 223–269.

Goldberg, P. K., & Pavcnik, N. (2007). Distributional effects of globalization in developing countries. *Journal of Economic Literature*, *45*(1), 39–82.

Grennan, M., Gupta, C., & Lederman, M. (2018). *Firm scope and spillovers from new product innovation: Evidence from medical devices* (Working Paper No. 25183). Cambridge, MA: National Bureau of Economic Research.

Grossman, G. M., & Helpman, E. (1991). *Innovation and growth in the global economy.* Cambridge, MA: MIT Press.

Halpern, L., Koren, M., & Szeidl, A. (2015). Imported inputs and productivity. *American Economic Review, 105*(12), 3660–3703.

Helpman, E., & Krugman, P. R. (1985). *Market structure and foreign trade: Increasing returns, imperfect competition, and the international economy.* Cambridge, MA: MIT Press.

Holmes, T. J., & Schmitz, J. A., Jr. (2001). A gain from trade: From unproductive to productive entrepreneurship. *Journal of Monetary Economics, 47*(2), 417–446.

Iacovone, L., & Javorcik, B. S. (2010). Multi□product exporters: Product churning, uncertainty, and export discoveries. *The Economic Journal, 120*(544), 481–499.

Javorcik, B. S. (2004). Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages. *American Economic Review, 94*(3), 605–627.

Khandelwal, A. (2010). The long and short (of) quality ladders. *The Review of Economic Studies, 77*(4), 1450–1476.

Khandelwal, A. K., Schott, P. K., & Wei, S.-J. (2013). Trade liberalization and embedded institutional reform: Evidence from Chinese exporters. *American Economic Review*, *103*(6), 2169–2195.

Kremer, M. (1993). The O-ring theory of economic development. *The Quarterly Journal of Economics*, *108*(3), 551–575. Kugler, M., & Verhoogen, E. (2012). Prices, plant size, and product quality. *The Review of Economic Studies*, *79*(1), 307–339.

Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, *70*(2), 317–341.

Lovo, S., & Varela, G. (2020). *Internationally linked firms, integration reforms, and productivity: Evidence from Pakistan* (Policy Research Working Paper No. 9349). Washington, DC: World Bank.

Ma, Y., Tang, H., & Zhang, Y. (2014). Factor intensity, product switching, and productivity: Evidence from Chinese exporters. *Journal of International Economics*, *92*(2), 349–362.

Manova, K., & Yu, Z. (2017). Multi-product firms and product quality. *Journal of International Economics*, *109*, 116–137.

Manova, K., & Zhang, Z. (2012). Export prices across firms and destinations. *The Quarterly Journal of Economics*, *127*(1), 379–436.

Melitz, M. J. (2003). The impact of trade on intra□industry reallocations and aggregate industryproductivity. *Econometrica*, *71*(6), 1695–1725.

Nishimizu, M., & Robinson, S. (1984). Trade policies and productivity change in semi industrialized countries. *Journal of Development Economics*, *16*(1–2), 177–206.

Olley, G. S., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica, 64*(6), 1263–1297.

Pavcnik, N. (2002). Trade liberalization, exit, and productivity improvements: Evidence from Chilean plants. *The Review of Economic Studies, 69*(1), 245–276.

Roberts, M. J., & Tybout, J. R. (1991). Size rationalization and trade exposure in developing countries. In R. E. Baldwin (Ed.), *Empirical studies of commercial policy* (pp. 169–200). Chicago, IL: University of Chicago Press.

Rodrik, D. (1992). Closing the productivity gap: Does trade liberalization really help? In G. K. Helleiner (Ed.), *Trade policy, industrialization, and development: New perspectives* (chap. 6).Oxford: Clarendon Press.

Romer, P. M. (1994). The origins of endogenous growth. *Journal of Economic Perspectives*, *8*(1), 3–22.

Schott, P. K. (2004). Across-product versus within-product specialization in international trade. *The Quarterly Journal of Economics*, *119*(2), 647–678.

Tirmazee, Z. S. (2016). *Relative wage variation and industry location in Punjab* (CREB Working Paper No. 02-16). Lahore: Lahore School of Economics.

Topalova, P., & Khandelwal, A. (2011). Trade liberalization and firm productivity: The case of India. *The Review of Economics and Statistics*, *93*(3), 995–1009.

Verhoogen, E. A. (2008). Trade, quality upgrading, and wage inequality in the Mexican manufacturing sector. *The Quarterly Journal of Economics*, *123*(2), 489–530.

Yu, M. (2015). Processing trade, tariff reductions, and firm productivity: Evidence from Chinesefirms. *The Economic Journal*, *125*(585), 943–988.

4. Paper III: Where to locate? The correlation between spatial proximity and location choice of new firms: The case of Pakistan

Section 4.1: Introduction

The concentration of industrial activity has been widely studied and agglomeration - defined as the industry-level benefits resulting from the presence of different economic units within the same geographical location — occurs widely across economies (see Krugman, 1995; Duranton and Puga, 2004). There are many well-known examples in developed economies, like the computer industry in Silicon Valley (Delgado, Porter and Stern, 2010), and in developing economies, like the surgical goods and sports goods industries in Sialkot, Pakistan (Atkin et al., 2017; Nadvi, 1999).

In this paper, we empirically estimate the relationship between agglomeration and the formation of new firms as well as their scale of operations using district-level data from Punjab, Pakistan. We test two hypotheses: First, do new firms in an industry choose to locate in an area where there is a similar industrial activity? Second, do new firms chose to locate in an area where there is a diverse industrial activity? We expect localization and urbanization to have differential impacts. We test this by analyzing the effects of agglomeration on the arrival and scale of operations of new firms at the district level in 2010, incorporating industrial controls and socioeconomic characteristics at the district level using a combination of firm-level data and household survey data. We also test the scale at which agglomeration forces (localization and urbanization) are likely to affect where new firms decide to locate.

On the whole, the empirical literature has found that firms benefit from localization, which refers to the benefits accruing to firms that choose to locate in a specific region within a specific industry; some of these benefits include input sharing (because of accessibility to suppliers and mutually enforced contracts), labor pooling (because of the availability of specialized labor) and knowledge spillovers (because of shared information about products, production process, innovations, existing and new technology, marketing agendas, and research and development (see Parr, 2002; Marshall, 1920). These benefits also lead to cost savings with one well-known example being the case of U.S. food processing industries (Cohen and Paul, 2003). The literature has also studied the benefits of urbanization, which refers to the benefits to firms from choosing to locate in an area where there is a wide range of diverse industrial activity. These benefits include the presence of diversified suppliers, specialized labor and suppliers, market mechanisms, transportation facilities, infrastructure, and community facilities, which make certain areas more attractive for new firms to enter (Parr, 2002). New entrepreneurial activity is likely to take place in areas of geographic concentration while new firms also enter when they anticipate a developed market, existing suppliers, and the availability of low-cost factors of production.

Regional economic characteristics and governmental policies have also been found to make a difference in the location decision of firms as location factors significantly affect the formation of new establishments in a particular region, including market demand, agglomeration, market conditions, and factor cost. Government policies can also attract new firms to a particular area through government spending on local infrastructure and the provision of direct assistance to firms.

A common thread running through much of the empirical literature is that most of the research studying where firms decide to locate has focused on developed economies: Glaeser and Kerr (2009) used U.S. data to find that labor and suppliers have a strong impact on entrepreneurship and are the main drivers of new firm formation; Delgado et al. (2010) and Rosenthal and Strange (2010) also used U.S. data to examine the relationship between agglomeration and both firm arrival and scale of operations in the manufacturing sector.

Otsuka (2008) used Japanese data to determine the impact of regional characteristics on new firm formation while Jofre-Monseny et al. (2014) used Spanish data to examine the factors affecting the location choice of new manufacturing firms. One of the few countries analyzed in the context of this research is Pakistan, though the research is relatively scant: Burki and Khan (2010) also found that industrial concentration is a prominent characteristic of all districts of Punjab but is more apparent in the urbanized areas of the province. Azhar and Adil (2016) show that district-level agglomeration is positively correlated with the efficiency of firms. Some of the recent work on Pakistan also includes Chaudhry, Haseeb, and Haroon, (2017).

This paper has four main contributions which makes it distinct from much of the existing literature (including Rosenthal and Strange, 2010). First, to the best of our knowledge, this is the first paper to analyze the impact of agglomeration economies (using localization and urbanization) on the location decision of firms and their scale of operations in a developing country context. Pakistan is an interesting example to study since specific goods (such as footballs and surgical instruments) from firms in Punjab have been found to be internationally competitive and these firms tend to be located in industrial clusters. So, this raises the question of whether new firms in a country like Pakistan prefer to locate in regions where existing similar industrial activity or diverse industrial activity is present. Second, we specifically aim to test at which level agglomeration economies are important. The existing literature (which includes Rosenthal and Strange, 2010) has mostly paid attention to localization and urbanization at the small scale while ignoring the importance of localization and urbanization at the medium and large scale. We believe that in addition to small-scale firms, medium and large-scale firms are also important for new firms in terms of vertical integration, subcontracting, the possibility of innovation, investment in research and development, exploring new markets for potential goods and entry barriers. Third, the literature has only analyzed the impact of agglomeration economies on new small-scale firms while we also analyze the impact of agglomeration economies on new large-scale and medium scale. Lastly, this study uses unique firm-level data which contains information on the population of firms rather than a sample of firms, which has been the basis of most of the previous analyses.

Our empirical results show that firms in Punjab significantly benefit from locating in agglomerated regions. Firms tend to locate in areas where firms from the same industry are prevalent. Also, they tend to locate in areas where there is also a diverse set of medium-sized firms. The scale of operations of new entrants is larger in areas where large or medium-scale firms belonging to the same industry are present. The scale of operations also tends to be larger in areas of medium-scale urbanization. Interestingly we find that the presence of large firms from different industries in an area tends to negatively affect new firm entry.

The paper is organized as follows. Section 4.2 presents our empirical specification and Section 4.3 describes the dataset used in our empirical estimations. Section 4.4 discusses the results, Section 4.5 discusses the policy implication of our results, and the conclusions are presented in Section 4.6.

Section 4.2: Empirical Specification

Based upon Rosenthal and Strange (2010), we analyze how agglomeration economies such as urbanization and localization affect new firm entry and their scale of operations by estimating the following model:

$$Arrival_{id} = A_{id} = \beta_0 + \beta_1 localization_{id} + \beta_2 urbanization_d$$

$$+ \beta_3 X_d + \beta_{4i} + \beta_{5sp} + \varepsilon_{a,id} \tag{4.1}$$

$$Scale \quad of \quad operations_{id} = E_{id} = \alpha_0 + \alpha_1 localization_{id}$$

$$+ \alpha_2 urbanization_d + \alpha_3 X_d + \alpha_{4i} + \alpha_{5sp} + \varepsilon_{e,id} \tag{4.2}$$

Equation (4.1) explains how new firm entry represented by A_{id} in industry i and district d, is affected by localization, urbanization, and the socioeconomic characteristics of the district with industry fixed effects and sub-provincial fixed effects. Equation (4.2) is an analogous equation that gives a firm's scale of operations at their time of entry (E_{id}). In the equations, ε_a and ε_e are standard error terms, β_{4i} and α_{4i} are industry fixed effects, β_{5sp} and α_{5sp} are sub-provincial fixed effects and X_d represents a vector of socioeconomic characteristics of a particular district.We use the 2010 Directory of Industries (DOI) for 2010 to define the entrance of new firms and the scale of operations and the DOI 2006 to define the agglomeration variables.

The first dependent variable is a new firm entry which is measured by the total number of newly-established firms in industry *i* and district *d*, where a new entrant is defined as a firm established in 2010[21]. The second dependent variable, *Scale of Operations* (E_{id}), is measured by the total employment of firms established in 2010 for an industry *i* and district *d* (as defined by Rosenthal and Strange, 2010).

We decompose agglomeration into localization and urbanization economies[22]. We use different measures of agglomeration economies considered in the literature, which are constructed for two different periods: 2010 and 2006. We measure localization economies using four different proxies: The first measure is the own-industry employment (adopted from Beaudry and Schiffauernova, 2009 and Rosenthal and Strange, 2010) and is constructed as the sum of employees in the industry *i* and district *d,* which allows us to examine how the presence of the same industry leads to the new firm formation in a specific area. To ease interpretation, we transform the variables into logarithmic form. But it is important to point out that authors like Beaudry and Schiffauernova (2009) have pointed out the limitations of own-industry employment since it might not control for area or region size; for this reason, we also use alternative measures. Second, we calculate the number of industrial plants in the industry *i* and district *d* (Beaudry and Schiffauernova, 2009). Third, we use the location quotient which is measured using the sum of employment in the industry *i* and district *d* divided by employment in district *d* (following Henderson et al., 1995; Combes and Gobillon, 2015)

[21] The data set does not distinguish between new firms or the movement of firms within an area which is one limitation of the dataset. So, we classify all firms whose year of establishment is 2010 as new firms.

[22] The detailed review on the pros and cons of these agglomeration variables has been discussed by Beaudry and Schiffauernova (2009) andCombes and Gobillon, (2015).

$$Locquo_{id} = \frac{Emp_{id}}{Emp_d}$$

The location quotient has been widely used since it captures the intensity and density of interaction among firms. Lastly, we use the Herfindahl index which is a measure of industry concentration constructed from the shares of industries within local employment (following Beaudry and Schiffauernova, 2009; Combes and Gobillon, 2015):

$$Herfindahl\ Index_i = \sum_i \left(\frac{Emp_{id}}{Emp_d}\right)^2$$

Jacobs (1969) popularized the idea that industrial diversity, or urbanization, can have a positive impact on firms because of the potential cross-fertilization of ideas and transmission of innovations between industries. This has been formalized by authors such as Duranton and Puga (2001), and many measures of urbanization have been proposed in the literature. We measure urbanization or diversity in multiple ways: First, urbanization is measured by the level of employment in existing establishments within a particular district (Rosenthal and Strange, 2010; Beaudry and Schiffauernova, 2009). This allows us to assess how the presence of different kinds of industries leads to the new firm formation in a specific district. Second, we use the share of other industry employment measured through total employment in region d as a ratio of total employment (Beaudry and Schiffauernova, 2009). Third, we measure urbanization through the number of active industries in a region (following Beaudry and Schiffauernova, 2009). Fourth, we use the Ellison and Glaeser index (following Ellison and Glaeser, 1997; Beaudry and Schiffauernova, 2009) which is measured as:

$$E - G\ index_i = \frac{G - (1 - \sum_d (Z_d)^2)\ H_i}{(1 - \sum_d (Z_d)^2)\ (1 - H_i)}$$

Where Z_d is the share of the districts j's total employment as a ratio of total employment, $H_i$is the Herfindahl index and G is the Gini coefficient, calculated as:

$$\sum_d (S_{id} - Z_d)^2)\ H_i$$

Lastly, we use the Krugman specialization index which measures an absence of diversity. The Krugman index is a measure of the distance between the distributions of industry shares in the location and at the global level (Krugman, 1991; Combes and Gobillon, 2015). A diversity index can be constructed as the log of 1 minus the Krugman specialization index, where the Krugman specialization index is as follows:

$$\text{Krugman specialization index}_d = \sum_i \left| \frac{Emp_{id}}{Emp_d} - \frac{Emp_i}{Emp} \right|$$

Localization and urbanization are initially incorporated as aggregate measures and then disaggregated into three levels of establishment: small, medium, and large. Or in other words, we first look at the impact of the total number of other firms in an area (both within the same industry and in other industries) on new firm entry and size. This is followed by separate analyses of how the number of small, medium, and large firms in an area (again in the same industry and in other industries) affects new firm entry and size. This disaggregation is done because the impact of small firms in an area on new firm entry and size can be significantly different than the impact of medium and large firms in an area. One reason for this (discussed by Rosenthal and Strange, 2010) is that smaller firms tend to support the entry of other smaller firms because they rely more on shared infrastructure and agglomeration economies. Another potential reason for this differential impact is that large firms may be able to erect formal and informal barriers to new firm entry. In our analysis we define small firms as those firms with fewer than 10 workers, medium-sized firms are defined as those firms employing between 10 and 49 workers, and large-sized firms are those employing 50 or more workers[23].

In order to account for the socioeconomic factors (X_d) that affect the entry of new firms, we incorporate Punjab district-level control variables. These include the average age of the population, the percentage of the male population, the average income, the unemployment rate, the percentage of the population with primary education, the percentage of the population with secondary education, and the percentage of the population with tertiary education.

We instrument current values of agglomeration in 2010 with their past values from 2006 which is commonly done in the literature (Rosenthal and Strange, 2010; Chaudhry, Haseeb and Haroon, 2017). We use past values as instruments since authors like Reed (2015) have found that instrumenting current values with lagged values are a better solution to simultaneity bias than using lagged values alone. We also incorporate industry and sub-provincial region fixed effects to account for specific industry and regional characteristics that might have an impact on new firm formation in a specific industry and district. Examples of industry-level heterogeneity include different barriers to entry, levels of innovation, and technological shifts. There are four sub-provincial regions in Punjab which are central Punjab, southern Punjab, northern Punjab, and

[23] In order to define small, medium and large sized firms, we follow the methodology of Rosenthal and Strange, (2010).

western Punjab and the reason for including dummies for these is that there may be common characteristics at the sub-provincial level.

Section 4.3: Data sources and data description

Our analysis focuses on the province of Punjab, Pakistan, and uses data from the Government of Punjab's Directory of Industries (DOI). We chose Punjab for the analysis since it is the Pakistani province with the highest population and because a significant proportion of the industrial base is located there which allows us to look at a new firm formation in a wide range of sectors across a broad socio-economic and geographical spectrum. Punjab is also the province containing the main industry clusters. The DOI is a firm-level dataset that includes more than 16,000 firms in a particular year and includes information on firms' years of establishment, employment levels, and districts collected by the Punjab Bureau of Statistics, Government of Punjab (for more information on the Punjab Bureau of Statistics, see http://www.bos.gop.pk/). We classify each firm in the DOI into industries using the Pakistan Standard Industrial Classification of All Economic Activities (PSIC) (Rev. 4) and the classification is in line with the International Standard Industrial Classification of All Economic Activities (ISIC) (Rev. 4) (For more details on PSIC classification see, http://www.pbs.gov.pk/sites/default/files/other/PSIC_2010.pdf). We have used the DOI to measure the new firms and their scale of operations in 2010 and local conditions (localization and urbanization) in two different periods 2010 and 2006. Socioeconomic characteristics at the district level are incorporated using the Government of Punjab's Multiple Indicator Cluster Survey (MICS).

We begin by presenting some basic summary statistics: Table 1 reports the total number of new establishments (arrival), total number of workers working in new firms (scale of operations), average localization, and average urbanization at aggregated and disaggregated levels. As the table shows 97 new firms entered the manufacturing sector in 2010, employing 3,323 employees. The table also shows that on average there are 494 workers working in small firms in the same industry within a district. Similarly, there are on average 1,408 and 2,525 workers working in medium and large-scale firms within the same industry. On average, 4,142 workers are working in small-scale firms in a district. Similarly, there are on average 11,637 and 28,293 workers in medium and large-scale firms in a district.

Section 4.4: Findings and Analysis

The results of our models (equations 1 and 2) are presented in Tables 2, 3, 4A, and 4B, all of which use lagged values as instruments for current values of agglomeration. Tables 2 and 3 reports the tobit estimates for the impact of agglomeration on the arrival of new firms and the marginal effects for the impact of agglomeration on the arrival of new firms respectively. Similarly, Tables 4A and 4B report the tobit estimates for the impact of agglomeration on the scale of operations of new firms and the marginal effects for the impact of agglomeration on the scale of operation of new firms respectively. Tables 5 and 6 report the estimates from a Heckman model for the impact of agglomeration on the arrival of new firms and the scale of operations of new firms respectively[24].

[24]Other researchershave used Poisson or negative binomial regressions to estimate similar models, however we use tobit estimations because of our limited sample size and greater incidence of zero values.

Table 7 reports the estimates for the arrival of new firms and the scale of operations of new firms using an alternative instrumental variable. Tables 8 and 9 report our results using panel data estimations. Table 10 presents estimates using disaggregated localization and urbanization measures and Table 11 presents estimates for small and medium-sized new firms.

Subsection 4.4.1: Impact of localization and urbanization on new firm arrival

We first analyze the correlation of localization and urbanization with the arrival and scale of new firms. Table 2 reports the tobit results and table 3 reports marginal effects for our models analyzing the impact of localization and urbanization at an aggregated level using various measures of localization and urbanization (following Beaudry and Schiffauernova, 2009; Henderson et al., 1995; Combes, 2000; Combes and Gobillon, 2015; Duranton and Puga, 2001; Krugman, 1991).

In tables 2 and 3, column (1) presents estimates measuring localization as own-industry employment for localization, column (2) measures localization using the number of industrial plants in an industry and region, column (3) measures localization using the location quotient and column (4) measures localization through the Herfindahl index with urbanization being measured using the level of employment in a region across all four specifications. We then used additional measures of urbanization including the share of other industry employment (presented in column 5), the number of active industries in a region (presented in column 6), the Ellison and Glaeser index (presented in column 7), and the diversity index (presented in column 8) and we control for localization across specifications (5) – (8).

In column (1) of Table 2, we find that localization (measured through own industry employment) is positively and significantly correlated with the entry of new firms. We also find that our results are robust to other measures of localization which include the number of industrial plants in an industry, the location quotient, and the Herfindahl index (presented in columns 2, 3, and 4). This suggests that firms are more likely to enter into regions that have similar industrial activity in an area for reasons such as the possibility of knowledge spillovers, input sharing, and labor pooling. Our results for localization (measured through own industry employment) are also consistent when we use different proxies for urbanization (presented in columns 5 to 8).

The results for urbanization show that new firms do not choose to enter into areas with higher urbanization (presented in columns 1 to 4). We also find that our results are consistent across various measures of urbanization, i.e. the share of other industry employment, the number of active industries in a region, the Ellison and Glaeser index, and the diversity index (presented in columns 5 to 8). This reinforces the idea that localization economies are more beneficial for manufacturing firms in Pakistan than urbanization economies. Our findings are broadly in line with the results found for other countries like the US and Japan (Helsley and Strange, 2002; Otsuka, 2008; Bosma et al., 2006; Figueiredo et al., 2009; Rosenthal and Strange, 2010) though this work found that both localization and urbanization economies have a significant impact on firm size and entry. But the literature has also found that urbanization economies may only have a positive effect in the case of knowledge-intensive and technologically advanced industries (Jofre-Monseny, Marín-López, and Viladecans-Marsal, 2014). Since industries in Pakistan tend to be less knowledge-intensive, it is possible that they fail to benefit from urbanization. Also, urbanization effects are expected to be more important in industries that employ workers whose skills are not industry-specific, and it is very possible that Pakistan firms are more inclined towards industry-specific skills.

Subsection 4.4.2: Impact of localization and urbanization on the scale of new firms

Our second model looks at the correlation of agglomeration on the size of new firms. Table 4A reports the tobit results and table 4B reports marginal effects for our models analyzing the impact of localization and urbanization on the scale of new firms in the manufacturing sector of Punjab.

The results show that localization has a positive relationship with the scale of operations of new firms (measured through own industry employment presented in column 1). Our results for localization economies are robust to other measures of localization which include the number of industrial plants in the industry, the location quotient, and the Herfindahl index (presented in columns 2, 3, and 4 respectively). Our results for localization (measured through own industry employment) are also consistent when we use different proxies for urbanization (presented in columns 5 to 8). We do find weak evidence of a negative relationship between urbanization economies and the scale of operation of new firms but that is not consistent across different specifications and choice of measure of urbanization (presented in columns 1 to 8). Similar to our findings of urbanization for new firms we also find the negative impact for the scale of operations which suggests that firms' scale of operation is likely to be smaller in presence of urbanization economies as there could be high competition for labor because of the presence of more industries and thus generating more demand for labor which could possibly deter firms to enter with a large scale of operation.

Subsection 4.4.3: Accounting for self-selection in the model of new firm firms and scale of operations

One potential issue that can occur when looking at the entry ofnew firms and the scale of new firms is the problem of selection bias which may arise if only a subset of industry-region combinations experiences entry of firms. In such case if our econometric analysis is restricted to this non-random subset of active industry-region combinations (which have experienced positive entry), then our estimates estimates would be biasedbecause of self-selection. One way to correct the selection bias by estimating a Heckman model comprising of a two-equation system (Heckman, 1974, 1978, 1979). The Heckman model allows us to separate the two decisions i.e., the decision to enter and then how many firms enter in each industry-region combination. In the first stage, we estimate the probability of a firm entering into a particular industry-district combination and in the second stage, we estimate the number of firms that enter as a function of agglomeration externalities and the inverse Mill's ratio calculated from the first step to control for selection bias.

The results from the second stage of the Heckman model for entry of new firms are presented in table 5 and the results for the scale of operation are presented in table 6. We find that the results that we obtain using different proxies of localization, which are own industry employment, the number of industrial plants in the industry, the location quotient, and the Herfindahl index (presented in columns 1 to 4 respectively), are consistent with our original results for new firms (presented in table 2 and 3) and scale of operation (presented in table 4A and 4B). We find that our results for localization economies remain significant across specifications and that we still find that urbanization is not positively correlated (though we did find some evidence of a negative effect which is not consistent across specifications) with the entry of new firms and scale of operation. We also find that the value of lambda in the Heckman estimates is insignificant across all specifications which suggests that the problem of selection is not playing a major role in our estimations.

Subsection 4.4.4: Testing the impact on new firm arrival and scale of operations using an alternate instrumental variable

In order to test the robustness of our results, we also use an alternate instrumental variable in our analysis. In particular, we use a natural experiment that has induced a sizeable, localized shock in some of the districts of Punjab in 1997 to instrument for agglomeration economies in 2006 and 2010. More specifically, we utilize the building of a major motorway (the M-2) as a natural experiment: The M-2 motorway was a 334 highway inaugurated in 1997 that was built between Islamabad and Lahore in order to connect the capital city with many major districts. It was the first major motorway built in South Asia and this also allowed many firms to relocate or new firms to enter the region on the route. We expect that the development of the motorway would have led many new firms to enter into, or existing firms to relocate to, the districts that it passed through. So while the construction of the motorway should be correlated with agglomeration, it should not be directly related to new firms in 2010; or in other words, we expect the presence of the M-2 to be a localized shock that affected some of the districts and created agglomeration economies which in turn affected new firms. The districts that the M-2 passes through are Lahore, Sheikhupura, Nankana Sahib, Hafizabad, Sargodha, Jhelum, Chakwal, and Rawalpindi. We create a dummy variable for the M-2 which equals to 1 for the above-mentioned districts and 0 otherwise. We then instrument agglomeration with M-2 to see the impact on new firms in 2010.

We present our results with the M-2 as an instrument for agglomeration in table 7. We find that our results for localization are consistent and robust to our earlier results in table 2. We also find that our urbanization is still not positively correlated with new firms (our results are inconsistent and switche sign). Our results also find that the partial R-squared of the first stage for our endogenous covariate and instrument are high enough to reject the hypothesis of a weak instrument. Finally, we also conduct a Sargan test of overidentifying restriction and the p-value suggests that we can reject the null hypothesis of valid over-identifying restrictions.

Subsection 4.4.5: Testing the impact on new firm arrival and scale of operations using panel data

In order to address the concern that certain important variables were omitted in our previous analyses, we then use panel data estimation techniques in order to control for unobserved characteristics and allow for a sufficient amount of time period in order to see the impact. We use 2016 and 2010 data to measure new firms, the scale of operation, localization, and urbanization, and we instrument localization and urbanization in 2016 with 2010 values and localization and urbanization in 2010 with 2006 values[25]. Table 8 present the marginal effects for our panel data estimates for the arrival of new firms and table 9 presents the marginal effect for our panel data estimates for the scale of operations of new firms. Consistent with our earlier findings, we find that localization has a

[25] The panel data would be at industry and district combinations which means industry- district combination in 2016 and 2010 has been used to create the panel.

significant and positive impact on the arrival of new firms (presented in columns 1 to 4 of table 8) and the scale of operation of new firms (presented in columns 1 to 4 of table 9) using various measures of localization and urbanization. However, we do not find any significant impact of urbanization on the arrival of new firms (presented in columns 5 to 8 of table 8) and the scale of operation of new firms (presented in columns 5 to 8 of table 9). So we can conclude that the results from the panel are consistent with those found in our previous models.

Subsection 4.4.6: Testing the impact of disaggregated levels of localization and urbanization on the arrival and scale of new firms

We present the tobit estimates and marginal effects for disaggregated levels of localization and urbanization for the arrival of new firms (presented in columns 1 and 2) and the scale of operations of new firms (presented in columns 3 and 4) in Table 10[26]. The results in columns (1) and (2) of Table 10 show that the presence of localization at all levels, small, medium, and large, leads to the greater entry of new firms or in other words the presence of the small, medium, and large-scale firms from the same industry increase firm entry in an area which reinforces the idea that new firms find benefits in locating near similar firms. It is also useful to compare the sizes of the estimated coefficients in the disaggregated analysis: the results suggest that the relationship between localization and firm entry is greater for small firms than for medium and large firms or the presence of small neighboring firms of the same industry has a greater impact on firm entry than the presence of large neighboring firms of the same industry. One reason for this could be that large firms tend to be more vertically integrated which

[26] Some of the existing work using disaggregated measures of localization includes Chaudhry, Haseeb, and Haroon, M. (2017).

entails fewer opportunities for smaller firms, which tend to be subcontractors, to enter a market.

The results of the impact of disaggregated urbanization on firm arrival in column (2) show that new firms tend to enter areas in which there are a greater number of medium-sized firms from all industries. This may be because new firms find it easier to initiate contracts with existing medium-sized employers (as opposed to larger firms who may be engaged in more long-term contracts) in an area.

The most interesting result (shown in column 1) is that the presence of large firms across all industries in an area has a negative impact on new firm entry which may be because larger firms are more attractive employers and so they get the best workers which make it difficult for new firms to attract the required number of workers regardless of their sector. Another potential reason for this result is that new firms might not choose to enter areas where large firms are operating because they anticipate that survival in the latter's presence may be difficult, given their lower-cost advantage especially if resources in the area are constrained.

The last set of findings are in contrast to those found by Rosenthal and Strange (2010) for the U.S. who found that only localization at medium scale attracts new firm entry while the other localization variables are insignificant while our results show that localization at all levels attracts new firms. Also, the earlier literature finds that urbanization at a small-scale fosters entry while we find that it is an urbanization at a medium scale that fosters entrepreneurship. This shows that agglomeration economies are important in developed and developing countries while the level at which it is fostering new firm's decisions vary across countries.

The disaggregated level analysis shows that localization is positively related and significant at the medium and large scale with the scale of operation of new firms (presented in columns 3 and 4). This means that the greater is the similar activity in an area the higher is the scale of operation of new entrants. It also shows that if there are a greater number of medium and large-scale firms of the same industry present in an area, then newer entrants tend to be larger. Or in other words, new firms tend to be larger in areas in which there are a greater number of medium-sized firms across industries and tend to be smaller in areas where there are a larger number of large-scale firms across industries. It may be that entrants need to be large enough to either survive or compete when the existing firms are bigger, and therefore there is a selection effect (dissuading small firms from entry) which goes along with the results from the entry regressions discussed above. But entrants tend to be smaller if there are a greater number of large firms in an area which may be because large firms in an area may erect formal and informal barriers to limit the entry of larger entrants. Another important factors could be market size which limits demand which in turn restricts the scale of new firms especially in domestic (non-export) sectors.

These results for the scale of operation model contrast with the earlier findings of Rosenthal and Strange (2010) who find that for the manufacturing industries urbanization at a small establishment-level increases scale of operation of entrants while we find that small firms have no significant impact. For localization, their findings show that it is the medium-sized establishments that have an impact which is similar to our findings as well.

This may show that for the scale of operations model, the localization economies in developing countries are similar to those of developed economies while the impact of urbanization economies may differ.

Subsection 4.4.7: Testing the impact of localization and urbanization on new firm entry by disaggregating arrival on the basis of scale of operation

In the previous sections, we analyzed whether localization and urbanization at an aggregated and disaggregated level attract all new firms. Table 11 shows the results of the impact of localization and urbanization on new small-scale firms and new medium-scale firms. We do not estimate the model for new large-scale firms due to the limited sample size. Columns 1 and 3 show the result of aggregated localization and urbanization on new firms depending on their scale of operation that is small and medium. Columns 2 and 4 show the impact of disaggregated localization and aggregated urbanization on a small scale and medium scale new firms.

The results show that localization at an aggregated level impact new small and medium scale firms which indicates that new small and medium scale firms are attracted to areas with similar activity, while these firms do not choose to locate in areas with the presence of diversified activity. The results in columns 2 and 4 show that it is the presence of small and medium scale firms from the similar industry that attract new small and medium scale firms. Localization at a large scale does not play a significant role in attracting new small and medium scale firms. This shows that small and medium-scale firms are prime drivers of new firms in an area. The results also suggest that large-scale firms from different industries negatively affect small and medium scale firms which indicates that the presence of large-scale firms deters new small and medium scale to enter an area.

Subsection 4.4.8: Analysis and Policy Discussion

Our study has important implications for economic development and public policy in the context of a developing country in particular, and for industries characterized as low-skilled or less technologically advanced in general. We have highlighted the mechanisms through which entrepreneurial activity can be enhanced in a developing country context and our results imply that firms are more likely to enter areas where there is already significant industrial concentration. We believe that this has promising implications for industrial policy aimed at building industrial zones and special economic zones in developing countries. Our study points out that industrial clusters or special economic zones that bring firms of similar groups of firms together hold greater potential than the development of clusters that contain more diverse firms.

This also has implications for government policy aimed at countering regional disparity in developing countries: First, it implies that governments may not be optimally utilizing resources when they give incentives to individual firms to go to underdeveloped areas where there are lower levels of preexisting industrial activity. Second, it suggests that there might be a need for governments to develop policies that attract a critical mass of firms to underdeveloped areas (through initiatives like industrial zones or free trade zones) before any significant industrial development can take place in these areas.

Section 4.5: Conclusion

This paper estimates the impact of agglomeration on new firms' formation and scale of operations in Punjab, Pakistan. While most of the existing literature has examined this relationship in developed countries like the U.S. and Japan, our results are novel in that they analyze this relationship in a developing country context. We use data from the Punjab Directory of Industries to measure how local conditions in an area (measured by localization and urbanization) affect the arrival and scale of operations of new firms in Punjab, Pakistan. In other words, our analysis has focused on whether new firms tend to locate in areas where the existing industrial activity is geographically concentrated.

Our findings have shown that the presence of the small, medium and large firms in one industry attracts new firms from the same industry in a particular district. Additionally, new firms are attracted to districts where there is diverse employment (employment in different industries) in medium-sized firms. Also, new firms tend to be larger when there are more medium and large-scale firms from the same industry already present in a district and new firms tend to be larger in districts where there are more medium-scale firms across industries. Finally, we find that the presence of large firms across all industries in a district has a negative impact on new firm entry and size.

These results imply that new firms enter agglomerated districts and that the local conditions of a district have a significant impact on new establishments and their scale. The results from our firm-level analyses are broadly consistent with the findings of earlier studies from other countries (see Otsuka, 2008; Rosenthal and Strange, 2010; Delgado et al., 2010; Bosma et al., 2006; Figueiredo et al., 2009).

However, the results from our disaggregated level analyses do differ from those found in earlier studies from developed countries: While the analyses for developed countries showed that only localization at medium scale attracts new firms, we find that new firms in Pakistan are more likely to locate in areas where there is localization at the small, medium and large scales which shows that localization at all levels may play a significant role in attracting new firms in developing countries. Also, while the analyses done using data from developed countries found that urbanization at the small scale attracted new firms, we find that it is the medium scale activity that attracts new firms as firms in the case of Pakistan. We also find that urbanization at a large scale is likely to deter the entry of new firms in Pakistan which implies that new firms do not locate enter into areas with preexisting large urbanization activity which might be the result of entry barriers erected by these large firms.

Table 4.1: Number of new establishments, the scale of operations, and average localization and urbanization at aggregated and disaggregated levels in Punjab

Total new establishments	97
Total workers at new establishments (scale of operations)	3,323
Average employment in own industry within district (localization)	
All size establishments	12944
Small establishments (< 10 workers)	494
Medium establishments (10–49 workers)	1408
Large establishments (50 or more workers)	2525
Average employment in all industries within district (urbanization)	
All size establishments	55227
Small establishments (< 10 workers)	4142
Medium establishments (10–49 workers)	11637
Large establishments (50 or more workers)	28293

Source: Authors' own calculations using Punjab Directory of Industries, Government of Punjab

Table 4.2: Tobit estimation for the correlation of agglomeration and new firm entry

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Localization								
Own-industry employment	4.092***				2.287***	2.667***	1.987***	2.375***
	(1.166)				(0.451)	(0.519)	(0.424)	(0.549)
Number of industry plants in industry and region		5.063***						
		(1.377)						
Location quotient			20.21***					
			(4.457)					
Herfindahl index				0.975**				
				(0.465)				
Urbanization								
Level of employment in a region	-10.86**	-9.337**	-6.287**	-7.369**				
	(5.003)	(4.278)	(2.933)	(3.109)				
Share of other industry employment					-166.9**			
					(70.51)			
Number of active industries in a region						-14.71***		
						(4.895)		
Ellison and Glaeser index							5.239	
							(15.33)	
Diversity Index								-0.0057
								(0.238)
N	443	443	397	397	443	397	443	397
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' own calculations.

Note: Dependent variable is new firm in 2010 and independent variables include localization and urbanization. Localization is measured through own-industry employment, number of industry plants in industry *i* and district *d,* location quotient and Herfindahl index. Urbanization is measured through level of employment in the existing establishments in a region, share of other industry employment, number of active industries in a region, Ellison and Glaeser index and diversity index. We take logs of all independent variables except indices. Controls include socio economic characteristics of a district such as average age of population, average income, average income squared, unemployment rate, percent population male, percent population with higher education, percent population with secondary education, percent population with primary education. Estimations also control for sub-provincial fixed effects. Standard errors are in parenthesis. N refers to the sample size. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4.3: Marginal effects from tobit estimation for the correlation of agglomeration and new firm entry

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Localization								
Own-industry employment	0.424**				0.407***	0.487***	0.376***	0.535***
	(0.182)				(0.0873)	(0.105)	(0.0884)	(0.133)
Number of industry plants in industry and region		0.487***						
		(0.186)						
Location quotient			1.510***					
			(0.412)					
Herfindahl index				0.284***				
				(0.0822)				
Urbanization								
Level of employment in a region	-0.182	-0.0825	-1.590***	-1.264***				
	(0.396)	(0.308)	(0.497)	(0.489)				
Share of other industry employment					-13.61**			
					(5.437)			
Number of active industries in a region						-1.274***		
						(0.395)		
Ellison and Glaeser index							0.00787	
							(1.104)	
Diversity Index								-0.0864***
								(0.0266)
N	443	443	397	397	443	397	443	397
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' own calculations.

Note: Dependent variable is new firm in 2010 and independent variables include localization and urbanization. Localization is measured through own-industry employment, number of industry plants in industry *i* and district *d*, location quotient and Herfindahl index. Urbanization is measured through level of employment in the existing establishments in a region, share of other industry employment, number of active industries in a region, Ellison and Glaeser index and diversity index. We take logs of all variables except indices. Controls include socio economic characteristics of a district such as average age of population, average income, average income squared, unemployment rate, percent population male, percent population with higher education, percent population with secondary education, percent population with primary education. Marginal effects are presented here computed at margins by treating independent variables as continuous rather than dummy variables and using margins command rather than mfx using STATA. Estimations also control for sub-provincial fixed effects. Standard errors are in parenthesis. N refers to the sample size. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4.4a: Tobit estimation for the correlation of agglomeration and scale of new firms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Localization								
Own-industry employment	173.7***				140.3***	182.6***	120.1***	162.6***
	(34.77)				(30.51)	(38.49)	(28.33)	(38.44)
Number of industry plants in industry and region		216.7***						
		(41.76)						
Location quotient			1,585***					
			(350.9)					
Herfindahl index				54.39*				
				(31.65)				
Urbanization								
Level of employment in a region	-755.2***	-691.8***	-390.5**	-448.4**				
	(199.1)	(186.5)	(196.5)	(196.1)				
Share of other industry employment					-10,904**			
					(4,871)			
Number of active industries in a region						-987.9***		
						(325.8)		
Ellison and Glaeser index							232.4	
							(984.5)	
Diversity Index								6.843
								(16.69)
N	422	422	397	397	422	397	422	397
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' own calculations.

Note: Dependent variable is a scale of a new firm in 2010 and the independent variables include localization and urbanization. Localization is measured through own-industry employment, a number of industrial plants in the industry i and district d, location quotient, and Herfindahl index. Urbanization is measured through the level of employment in the existing establishments in a region, the share of other industry employment, the number of active industries in a region, the Ellison and Glaeser index, and the diversity index. We take logs of all independent variables except indices. Controls include socio-economic characteristics of a district such as average age of the population, average income, average income squared, unemployment rate, percent population male, percent population with higher education, percent population with secondary education, percent population with primary education. Estimations also control for sub-provincial fixed effects. Standard errors are in parenthesis. N refers to the sample size. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4.4b: Marginal effects from tobit estimation for the correlation of agglomeration and scale of new firms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Localization								
Own-industry employment	23.68***				23.38***	28.57***	22.50***	29.06***
	(5.431)				(5.265)	(6.751)	(5.495)	(7.553)
Number of industry plants in industry and region		21.96***						
		(5.048)						
Location quotient			82.23***					
			(27.08)					
Herfindahl index				12.28***				
				(4.211)				
Urbanization								
Level of employment in a region	-24.39*	-17.41	-286.8**	-104.8**				
	(13.55)	(12.09)	(97.36)	(41.53)				
Share of other industry employment					-660.5**			
					(317.3)			
Number of active industries in a region						-89.03**		
						(26.19)		
Ellison and Glaeser index							8.897	
							(62.42)	
Diversity Index								-2.081*
								(1.107)
N	422	422	397	397	422	397	422	397
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' own calculations.

Note: Dependent variable is a scale of a new firm in 2010 and independent variables include localization and urbanization. Localization is measured through own-industry employment, a number of industrial plants in the industry i and district d, location quotient, and Herfindahl index. Urbanization is measured through the level of employment in the existing establishments in a region, the share of other industry employment, the number of active industries in a region, the Ellison and Glaeser index, and the diversity index. We take logs of all variables except indices. Controls include socio-economic characteristics of a district such as average age of the population, average income, average income squared, unemployment rate, percent population male, percent population with higher education, percent population with secondary education, percent population with primary education. Marginal effects are presented here computed at margins by treating independent variables as continuous rather than dummy variables and using margins command rather than mfx using STATA. Estimations also control for sub-provincial fixed effects. Standard errors are in parenthesis. N refers to the sample size. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4.5: Heckman model: estimates for the correlation of agglomeration and new firm entry (second stage results)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Localization								
Own-industry employment	1.100**				1.223***	1.732***	1.104**	1.753***
	(0.444)				(0.449)	(0.575)	(0.437)	(0.560)
Number of industry plants in industry and region		1.257**						
		(0.586)						
Location quotient			9.569**					
			(4.703)					
Herfindahl index				0.988**				
				(0.468)				
Urbanization								
Level of employment in a region	-0.481	-0.380	1.326*	0.203				
	(0.625)	(0.643)	(0.801)	(0.761)				
Share of other industry employment					-10.16			
					(21.24)			
number of active industries in a region						-0.896		
						(1.379)		
Ellison and Glaeser index							-1.385	
							(12.06)	
Diversity Index								0.280
								(0.210)
N	443	443	397	397	443	397	443	397
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' own calculations.

Note: Dependent variable is a scale of a new firm in 2010 and the independent variables include localization and urbanization. Localization is measured through own-industry employment, a number of industrial plants in the industry i and district d, location quotient, and Herfindahl index. Urbanization is measured through the level of employment in the existing establishments in a region, the share of other industry employment, the number of active industries in a region, the Ellison and Glaeser index, and the diversity index. We take logs of all independent variables except indices. Controls include socio-economic characteristics of a district such as average age of the population, average income, average income squared, unemployment rate, percent population male, percent population with higher education, percent population with secondary education, percent population with primary education. Estimations also control for sub-provincial fixed effects. Standard errors are in parenthesis. N refers to the sample size. *$p < 0.1$, **$p < 0.05$, *** $p < 0.01$.

Table 4.6: Heckman model: estimates for the correlation of agglomeration and scale of operation of new firms (second stage results)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Localization								
Own-industry employment	69.01*				77.15**	131.5***	76.25**	132.3***
	(37.54)				(37.79)	(48.80)	(37.46)	(47.67)
Number of industry plants in industry and region		77.53*						
		(37.10)						
Location quotient			842.7**					
			(383.0)					
Herfindahl index				16.34*				
				(7.67)				
Urbanization								
Level of employment in a region	-8.415	0.755	-135.5**	38.59				
	(53.43)	(54.32)	(66.82)	(65.01)				
Share of other industry employment					190.5			
					(1,807)			
Number of active industries in a region						-36.08		
						(115.5)		
Ellison and Glaeser index							-30.77	
							(1,015)	
Diversity Index								19.55
								(17.88)
N	443	443	397	397	443	397	443	397
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' own calculations.

Note: Dependent variable is a scale of a new firm in 2010 and the independent variables include localization and urbanization. Localization is measured through own-industry employment, a number of industrial plants in the industry i and district d, location quotient, and Herfindahl index. Urbanization is measured through the level of employment in the existing establishments in a region, the share of other industry employment, the number of active industries in a region, the Ellison and Glaeser index, and the diversity index. We take logs of all independent variables except indices. Controls include socio-economic characteristics of a district such as average age of the population, average income, average income squared, unemployment rate, percent population male, percent population with higher education, percent population with secondary education, percent population with primary education. Estimations also control for sub-provincial fixed effects. Standard errors are in parenthesis. N refers to the sample size. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4.7: Estimates using alternate instrumental variable to estimate the correlation between agglomeration and new firm

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Localization								
Own-industry employment	0.126***				0.116***	0.138***	0.110***	0.156***
	(0.0355)				(0.0289)	(0.0341)	(0.0284)	(0.0402)
Number of industry plants in industry and region		0.253***						
		(0.0489)						
Location quotient			2.411***					
			(0.479)					
Herfindahl index				0.0962*				
				(0.0451)				
Urbanization								
Level of employment in a region	-0.0552	-0.0442	0.0104	-0.0752				
	(0.0370)	(0.0333)	(0.102)	(0.104)				
Share of other industry employment					-2.250			
					(3.017)			
Number of active industries in a region						-0.536*		
						(0.289)		
Ellison and Glaeser index							0.829	
							(1.312)	
Diversity Index								0.0281
								(0.0294)
Partial R-sq (first stage)	0.802	0.9283	0.9349	0.9763	0.8363	0.8246	0.5457	0.9891
F statistics (first stage)	872.688	2790.9	2764.39	7928.5	1100.71	905.058	258.813	17405.1
P-value of F-statistics (first stage)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Sargan statistics	0.292	0.2248	0.1879	0.3677	0.4936	0.0245	0.2896	0.0380
P-value of Sargan test	0.589	0.6354	0.6646	0.5443	0.4823	0.8754	0.5905	0.8453

N	443	443	397	397	443	397	443	397
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' own calculations.

Note: Dependent variable is the new firm in 2010 and the independent variables include localization and urbanization. Localization is measured through own-industry employment, a number of industrial plants in the industry *i* and district *d*, location quotient, and Herfindahl index. Urbanization is measured through the level of employment in the existing establishments in a region, the share of other industry employment, the number of active industries in a region, the Ellison and Glaeser index, and the diversity index. We take logs of all independent variables except indices. Controls include socio-economic characteristics of a district such as average age of the population, average income, average income squared, unemployment rate, percent population male, percent population with higher education, percent population with secondary education, percent population with primary education. We use an instrumental variable approach to estimate the model. Standard errors are in parenthesis. N refers to the sample size. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4.8: Marginal effects for the correlation of agglomeration and new firm entry using panel data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Localization								
Own-industry employment	0.0971***				0.101***	0.0937***	0.110*	0.117**
	(0.0328)				(0.0373)	(0.0344)	(0.0619)	(0.0490)
Number of industry plants in industry and region		0.0016**						
		(0.000632)						
Location quotient			0.997*					
			(0.544)					
Herfindahl index				0.104**				
				(0.0471)				
Urbanization								
Level of employment in a region	-0.0211	-0.0344	0.0433	0.0162				
	(0.259)	(0.241)	(0.218)	(0.228)				
Share of other industry employment					1.632			
					(1.967)			
Number of active industries in a region						0.0444		
						(0.346)		
Ellison and Glaeser index							0.137	
							(3.643)	
Diversity Index								0.0123
								(0.0149)
N	876	876	876	876	876	876	876	876
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' own calculations.

Note: Dependent variable is the new firm in 2010 and the independent variables include localization and urbanization. Localization is measured through own-industry employment, a number of industrial plants in the industry i and district d, location quotient, and

Herfindahl index. Urbanization is measured through the level of employment in the existing establishments in a region, the share of other industry employment, the number of active industries in a region, the Ellison and Glaeser index, and the diversity index. We take logs of all independent variables except indices. Controls include socio-economic characteristics of a district such as average age of the population, average income, average income squared, unemployment rate, percent population male, percent population with higher education, percent population with secondary education, percent population with primary education. Marginal effects are presented here computed at margins by treating independent variables as continuous rather than dummy variables and using margins command rather than mfx using STATA. Standard errors are in parenthesis. N refers to the sample size. *$p < 0.1$, **$p < 0.05$, *** $p < 0.01$.

Table 4.9: Marginal effects for the correlation of agglomeration and scale of new firms using panel data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Localization								
Own-industry employment	3.968***				4.234**	3.862***	4.750*	5.206**
	(1.420)				(1.662)	(1.474)	(2.611)	(2.226)
Number of industry plants in industry and region		0.0351						
		(0.0247)						
Location quotient			39.31*					
			(23.15)					
Herfindahl index				3.550**				
				(1.795)				
Urbanization								
Level of employment in a region	-0.682	-1.537	2.233	0.699				
	(11.64)	(9.021)	(9.528)	(9.976)				
Share of other industry employment					85.38			
					(91.92)			
Number of active industries in a region						1.884		
						(15.51)		
Ellison and Glaeser index							14.80	
							(171.4)	
Diversity Index								1.025
								(0.755)
N	876	876	876	876	876	876	876	876
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' own calculations.

Note: Dependent variable is a scale of a new firm in 2010 and the independent variables include localization and urbanization. Localization is measured through own-industry employment, a number of industrial plants in the industry i and district d, location quotient, and Herfindahl index. Urbanization is measured through the level of employment in the existing establishments in a region, the share of other industry employment, the number of active industries in a region, the Ellison and Glaeser index, and the diversity index. We take logs of all independent variables except indices. Controls include socio-economic characteristics of a district such as average age of the population, average income, average income squared, unemployment rate, percent population male, percent population with higher education, percent population with secondary education, percent population with primary education. Marginal effects are presented here computed at margins by treating independent variables as continuous rather than dummy variables and using margins command rather than mfx using STATA. Standard errors are in parenthesis. N refers to the sample size. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4.10: Tobit estimates and marginal effects for the correlation of agglomeration disaggregated into small, medium and large with new firm and scale of new firms

	(1)	(2)	(3)	(4)
	Tobit Estimates New firms	Marginal Effects New firms	Tobit Estimates Scale of new firms	Marginal Effects Scale of new firms
Localization				
Own-industry employment small	6.233***	0.410**	102.3	5.297
	(1.842)	(0.134)	(152.8)	(9.380)
Own-industry employment medium	2.605**	0.223**	116.5**	12.37**
	(1.282)	(0.107)	(54.01)	(5.892)
Own-industry employment large	2.468*	0.343*	416.0**	9.934*
	(1.214)	(0.171)	(119.2)	(4.699)
Urbanization				
Level of employment in a region small	-9.011	-0.0847	306.6	-5.920
	(6.723)	(0.192)	(201.1)	(12.38)
Level of employment in a region medium	5.092**	0.106*	403.7***	7.756*
	(2.025)	(0.045)	(144.3)	(3.767)
Level of employment in a region large	-4.779**	-0.162**	-350.3**	-4.348**
	(2.371)	(0.076)	(160.7)	(2.258)
N	443	443	422	422
Controls	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes

Source: Authors' own calculations.

Note: Dependent variable is the new firm in 2010 and the scale of operation of new firms, independent variables include localization and urbanization. Localization and urbanization are disaggregated into 3 categories small, medium, and large where small is defined as those firms with fewer than 10 workers, medium as those firms employing between 10 and 49 workers, and large establishments are those employing 50 or more workers. We take logs of all independent variables. Controls include socio-economic characteristics of a district such as average age of the population, average income, average income squared, unemployment rate, percent population male, percent population with higher education, percent population with secondary education, percent population with primary education.

Marginal effects are presented here computed at margins by treating independent variables as continuous rather than dummy variables and using margins command rather than mfx using STATA. Estimations also control for sub-provincial fixed effects. Standard errors are in parenthesis. N refers to the sample size. $*p < 0.1$, $**p < 0.05$, **

Table 4.11: Tobit estimates: Impact of agglomeration on firm arrival for manufacturing industry in Punjab

	(1)	(2)	(3)	(4)
	Small New Firms		Medium New Firms	
Localization				
Own-industry employment	3.680***		2.273***	
	(1.248)		(0.880)	
Own-industry employment small		3.071***		1.885**
		(0.985)		(0.775)
Own-industry employment medium		2.309***		1.915***
		(0.687)		(0.625)
Own-industry employment large		-0.574		-0.0410
		(2.876)		(0.871)
Urbanization				
Level of employment in a region	-10.11*		-5.643*	
	(5.220)		(2.535)	
Level of employment in a region small		-1.351		-0.693
		(0.857)		(0.647)
Level of employment in a region medium		-1.072		0.154
		(0.732)		(0.677)
Level of employment in a region large		-1.772*		-0.950*
		(0.721)		(0.573)
N	443	443	443	443
Controls	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes

Note: Dependent variable is small new firms in 2010 and medium new firms in 2010, independent variables include localization and urbanization. Localization and urbanization are disaggregated into 3 categories small, medium, and large where small is defined as those firms with fewer than 10 workers, medium as those firms employing between 10 and 49 workers, and large establishments are those employing 50 or more workers. We take logs of all independent variables. Controls include socio-economic characteristics of a district such as average age of the population, average income, average income squared, unemployment rate, percent population male, percent population with higher education, percent population with secondary education, percent population with primary education. Marginal effects are presented here computed at margins by treating independent variables as continuous rather than dummy variables and using margins command rather than mfx using STATA. Estimations also control for sub-provincial fixed effects. Standard errors are in parenthesis. N refers to the sample size. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

References

Azhar, A., & Adil, S. (2016). Effect of Agglomeration on Socio-economic Outcomes: A District Level Panel Study of Punjab. *The Pakistan Development Review*, 155-170.

Atkin, D., Chaudhry, A., Chaudry, S., Khandelwal, A. K., & Verhoogen, E. (2017). Organizationalbarriers to technology adoption: Evidence from soccer-ball producers in Pakistan. *The Quarterly Journal of Economics*, *132*(3), 1101-1164.

Burki A A, Khan M A (2010) Spatial Inequality and Geographic Concentration of Manufacturing Industries in Pakistan, in *Proceedings of 27th Annual General Meeting and Conference of the Pakistan Society of Development Economists,* Islamabad, Pakistan.

Beaudry, C., & Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research policy*, *38*(2), 318-337.

Chaudhry, T., Haseeb, M., & Haroon, M. (2017). Economic geography and misallocation in Pakistan's manufacturing hub. *The Annals of Regional Science*, *59*(1), 189 -208.

Cohen J P, Paul C J M (2003) Spatial and supply/demand agglomeration economies: State-and industry linkages in the US food system. *Empirical Economics*, *28*(4), 733-751. doi:10.1007/s00181-003 0156-6.

Combes, P. P., & Gobillon, L. (2015). The empirics of agglomeration economies. In *Handbook of regional and urban economics* (Vol. 5, pp. 247-348). Elsevier.

Delgado, M., Porter, M. E., & Stern, S. (2010). Clusters and entrepreneurship. *Journal of Economic Geography*, *10*(4), 495-518.

Duranton, G., & Puga, D. (2001). Nursery cities: urban diversity, process innovation, and the life cycle of products. *American Economic Review*, *91*(5), 1454-1477.

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In *Handbook of regional and urban economics* (Vol. 4, pp. 2063-2117). Elsevier.

Ellison, G., & Glaeser, E. L. (1997). Geographic concentration in US manufacturing industries: a dartboard approach. *Journal of political economy*, *105*(5), 889-927.

Figueiredo, O., Guimaraes, P., & Woodward, D. (2009). Localization economies and establishment size: was Marshall right after all?. *Journal of Economic Geography*, *9*(6), 853-868.

Glaeser, E. L., & Kerr, W. R. (2009). Local industrial conditions and entrepreneurship: how much of the spatial distribution can we explain?. *Journal of Economics & Management Strategy*, *18*(3), 623-663.

Henderson, V., Kuncoro, A., & Turner, M. (1995). Industrial development in cities. *Journal of Political Economy*, *103*(5), 1067-1090.

Jofre□Monseny, J., Marín□López, R., & Viladecans□Marsal, E. (2014). The determinants of localization and urbanization economies: Evidence from the location of new firms in Spain. *Journal of Regional Science*, *54*(2), 313-337.

Jacobs, J. (1969). The Economics of Cities Random House. *New York, NY.*

Krugman, P. R. (1997). *Development, geography, and economic theory* (Vol. 6). MIT press.

Marshall, A. (1920). *Principles of economics: an introductory volume*. Macmillan.

Nadvi, K. (1999). Shifting ties: social networks in the surgical instrument cluster of Sialkot, Pakistan. *Development and Change*, *30*(1), 141-175.

Parr, J. B. (2002). Missing elements in the analysis of agglomeration economies. *International Regional Science Review*, *25*(2), 151-168.

Rosenthal, S. S., & Strange, W. C. (2010). Small establishments/big effects: Agglomeration, industrial organization and entrepreneurship. In *Agglomeration economics* (pp. 277-302). University of Chicago Press.

5. Conclusion

There is significant evidence inthe existing literature that industrial clusters reduce costs and increase productivity, but the literature also suggests that agglomeration may result in firms' higher markups. Much of the existing empirical evidence on the impact of agglomeration has been based on data from developed economies. However, the firms in developing countries face additional constraints such as weak formal institutions, less efficient markets, costly research and development, weak contract enforcement, and binding financial constraints. The presence of agglomeration economies is more likely to be supportive and operative in presence of the constraints and help firms collaborate and act collectively. In the first part of this thesis, we test the impact of agglomeration economies in a developing country context using newly constructed firm-level panel data from Pakistan. Our findings suggest that, when looking at all the clusters, we do not find consistent evidence of agglomeration externalities in improving the productivity of firms though we do find that firms in agglomerated regions have higher costs and these firms can charge higher prices. We then exploit the fact that the Pakistani industrial sector is characterized by older and newer clusters to test if the impact of agglomeration is different for firms in older clusters as compared to firms in newer clusters. We find that the productivity of firms in the older clusters is negatively affected by agglomeration while the markups and costs of firms in the newer clusters are positively affected by agglomeration. The results seem to imply that a more nuanced approach to understanding the impact of agglomeration is required especially in countries that are characterized by both older and newer clusters.

Substantial gains in aggregate output arise when policy reforms, such as international trade liberalization, or changes in market fundamentals induce a reallocation from low- to high-performance firms within industries. There is also reallocation of output within a multiproduct firm through changes in product mix in response to changes in an economic environment. Drawing on existing extensive theoretical and empirical literature on multiproduct firms which builds on theories of industry dynamics by modeling endogenous product selection of firms in response to trade shocks we look at the various channels through which bilateral trade agreement affects firms in the context of the 2006 Pakistan-China Free Trade Agreement in the second part of this thesis. In particular, we have tested to see how lower Chinese tariffs on Pakistani goods and lower Pakistani tariffs on Chinese goods affected the products and productivity of firms.

We find that Pakistani firms added more products after the free trade agreement as a result of lower Chinese tariffs, and added fewer products because of lower Pakistani tariffs on imported Chinese inputs. We also find that firms improved product quality lowered prices, and increased the quality of their inputs as a result of lower Pakistan tariffs on imported Chinese inputs. Finally, we find that reductions in Chinese tariffs reduced firm-level productivity, while lower Pakistan tariffs increased productivity. The chapter illustrates the impact of bilateral trade agreements in a developing country context, and additionally, it also shows how a developing country is affected by a trade agreement with China. As China continues to cultivate economic ties with developing countries, it is important to gauge the economic impact of wide-ranging trade agreements on firms in these countries.

The formation of new firms is an important determinant of regional economic development and the industrial organization literature highlights agglomeration as one of the main factors affecting the formation and scale of operations of new firms. The third part of this thesis evaluates the impact of agglomeration in attracting new firms' in an area in Punjab, Pakistan. While most of the existing literature has examined this relationship in developed countries like the U.S. and Japan, our results are novel in that they analyze this relationship in a developing country context. We use data from the Punjab Directory of Industries to measure how local conditions in an area (measured by localization and urbanization) affect the arrival and scale of operations of new firms in Punjab, Pakistan. In other words, our analysis has focused on whether new firms tend to locate in areas where the existing industrial activity is geographically concentrated. Our findings suggest that new firms enter in agglomerated districts and that the local conditions of a district have a significant impact on new establishments and their scale.

This thesis has important implications for economic development and public policy in the context of a developing country in particular, and for industries characterized as low-skilled or less technologically advanced in general. We have highlighted the mechanisms through which entrepreneurial activity and firm's productivity and markups can be enhanced in a developing country context and our results imply that firms are more likely to enter areas where there is already significant industrial concentration. Our results also imply that firms in concentrated regions can collaborate or collude to act jointly. We believe that this has promising implications for industrial policy aimed at building industrial zones and special economic zones in developing countries.

This thesis also has implications for government policy aimed at countering regional disparity in developing countries: First, our results imply that governments may not be optimally utilizing resources when they give incentives to individual firms to go to underdeveloped areas where there are lower levels of preexisting industrial activity. Second, it suggests that there might be a need for governments to develop policies that attract a critical mass of firms to underdeveloped areas (through initiatives like industrial zones or free trade zones) before any significant industrial development can take place in these areas.