WIDER Working Paper 2014/065

Clustering, competition, and spillover effects

Evidence from Cambodia

Sokty Chhair and Carol Newman

March 2014

Abstract: The potential benefits of the geographical clustering of economic activity have been well documented in the literature, yet there is little empirical evidence quantifying these effects in developing country contexts. This is surprising given the emphasis in industrial policy on productivity growth and the potential gains that could be made by facilitating cluster formation. It is also possible that for some firms there may be disadvantages associated with locating close to competitors, in particular if they sell to customers located in the same geographic area. These represent a large proportion of firms in developing country settings at the early stage of industrialization, where physical infrastructure is underdeveloped and there are a large number of informal and service sector firms that often exclusively rely on customers in local markets. Using data on the population of all firms in Cambodia we investigate the pattern of firm clustering and explore the extent to which it leads to productivity-enhancing effects. We focus on two channels, a competition and a spillover channel and investigate the types of firms that benefit or suffer as a result of geographical clustering. We find strong negative competition effects associated with clustering for formal and manufacturing firms. We find some evidence of productivity spillovers for informal firms and firms in the manufacturing sector but they are not of a large enough magnitude to outweigh the negative competition effects observed.

Keywords: clustering, productivity spillovers, competition effects, informal firms, Cambodia
JEL classification: D2, L2, L6, L8, O1

Acknowledgements: This paper is one of a series of studies on industrial development in Africa produced in collaboration between the Africa Growth Initiative at Brookings, the African Development Bank (AfDB), and UNU-WIDER under their joint project 'Learning to Compete' (L2C). Our thanks go to John Page, John Rand, Måns Söderbom, Finn Tarp, and to participants at the UNU-WIDER L2C conference on Industrial Development and Policy for comments on an earlier draft of this paper. All errors are our own.

Tables and figures appear at the end of the paper.

[1]Trinity College Dublin; [2]Cambodian Economic Association; corresponding author: cnewman@tcd.ie

This study has been prepared within the UNU-WIDER project 'Learning to Compete: Accelerating Industrial Development in Africa', directed by John Page.

ISSN 1798-7237 ISBN 978-92-9230-786-8 https://doi.org/10.35188/UNU-WIDER/2014/786-8

Typescript prepared by Lisa Winkler at UNU-WIDER.

UNU-WIDER gratefully acknowledges the financial contributions to the research programme from the governments of Denmark, Finland, Sweden, and the United Kingdom.

1 Introduction

The geographical clustering of economic activity can impact on the productivity of firms in a number of different ways. It is expected that firms in a cluster can benefit from productivity improvements due to reduced transport costs, access to a common pool of labour (skilled or unskilled), technology spillovers and increased competitive pressure (for seminal work see Fujita et al. (1999); Krugman 1991; Marshall 1920). While there is a body of evidence on the existence of agglomeration economies in the developed world (see Duranton and Overman (2005) for evidence from the UK, and Audretsch and Feldman (1996) and Ellison and Glaeser (1999) for evidence from the USA), there is only a small, albeit growing, pool of evidence on the prevalence and benefits of clustering in developing country contexts.[1] This is surprising given the emphasis in industrial policy in developing countries on productivity growth and the potential gains that could be achieved through the agglomeration of firms.[2]

While the focus of most empirical work is on identifying the productivity *gains* from clustering, it is also possible that there are disadvantages associated with the clustering of firms, particularly for firms that compete to sell to customers located in the same geographic area. Increased competitive pressures will drive down the price of goods and services and erode mark-ups. Only the most productive firms will survive and many firms will be forced to locate far away from each other as a result. The extent to which productivity-enhancing or profit-eroding effects dominate will depend to a large extent on the types of firms, sectors, and context in question and thus is an empirical question worthy of investigation.

In developing country settings firms potentially have much to gain from clustering. Given that they are generally starting from a lower technological base, spillovers of new technologies and innovations are likely to significantly impact on their productivity, growth and survival probability. At the same time, competitive pressures are likely to be more pronounced in developing countries that are at the early stages of industrialization, at least in certain sectors, where physical infrastructure is underdeveloped leading producers to exclusively rely on customers in local markets.[3]

Differences in the types of firms might also make the nature and impact of agglomeration different to that observed in developed country settings. Many firms are informal or unregistered enterprises and as such face a different set of constraints in making decisions about where and what to produce.[4] Due to their small size, informal firms are expected to cluster (Rosenthal and Strange 2004), however, as pointed out by Moreno-Monroy (2012) their rationale for clustering might be very different to formal firms. It is more likely that their reasons for clustering are demand-driven; their small size means that they need to locate near their customer base. As such, competition effects may outweigh any positive productivity spillovers for these firms. There are also, however, possibilities for productivity spillovers resulting from agglomeration for informal

[1] Some exceptions, amongst others, include Fan and Scott (2003), Howard et al. (2011), Siba et al. (2012) and Fafchamps and Söderbom (2014).

[2] See, for example, Cortright (2006) who highlights the importance of understanding and promoting clusters for improving the performance of regional economies.

[3] Siba et al. (2012) find evidence of such competitive pressures, alongside positive productivity externalities, for manufacturing firms in Ethiopia. Overall they find that the net effect of clustering is close to zero.

[4] See Moreno-Monroy (2012) for a critical commentary on the treatment of informal firms in the literature on agglomeration economies in developing countries.

firms. For example, Overman and Venables (2005) highlight the potential for agglomeration economies in networks of small informal firms due to productivity spillovers from a concentration of labour or knowledge spillovers from formal to informal firms that interact through the supply chain. There is, however, very little empirical evidence available on the impact of clustering on informal firms, a gap that this paper aims to fill.

Another feature that distinguishes developing countries is that service sector firms make up a large proportion of economic activity. Theoretically service sector firms could benefit from similar agglomeration economies to manufacturing firms, such as, technology spillovers, access to a common pool of labour, or close proximity to input and output markets. Service firms might also cluster to have access to a centralized pool of customers. These firms are, however, likely to experience greater competitive pressures from clustering than manufacturing firms given that consumption of the service must take place at the point of sale limiting their options for selling to customers outside of the locality. While there is limited evidence on the extent and impact of the clustering of manufacturing firms in developing countries there is, to our knowledge, no empirical evidence on the impact of clustering on service sector firms.

In this paper we investigate the pattern of firm clustering in the setting of Cambodia and explore the extent to which it leads to productivity gains for different types of firms in different sectors. We consider both competition and technology spillover channels in explaining the pattern of clustering observed. We use data on the population of all firms (manufacturing and service sector firms in the formal and informal sector) in Cambodia for 2009 and 2011. The latter contains extensive information on the population of all establishments and is the main data source used in our analysis. The former only contains information on the location of the firm and the numbers employed. We use this information to capture changes in the characteristics of firm clusters over time. In addition, we match data from the 2008 Population Census to the Enterprise Census in order to incorporate control variables on the population density and the quality of local infrastructure into our analysis.

Cambodia is a particularly interesting case for exploring these effects given that infrastructure is still poorly developed and transport costs are high leading to a significant amount of clustering of economic activity. Moreover, most firms are informal or service sector firms. Tis paper addresses four main questions. First, we consider whether firms are more or less productive where there is greater clustering of economic activities. Second, we analyse whether different types of firms are impacted differently by the clustering of economic activities. Third, we explore whether there are productivity spillovers associated with clustering. Finally, we consider whether different types of firms are affected differently by productivity spillovers. In considering firm heterogeneity in the impact of clustering and the existence of spillovers we pay particular attention to differences between formal and informal firms, and manufacturing and service sector firms.

The rest of the paper is organized as follows. Section 2 describes the mechanisms through which clustering could impact on productivity at the firm level while Section 3 outlines our empirical approach to identifying these effects. Section 4 presents the data and describes the pattern of clustering of firms observed in the Cambodian case. Section 5 presents the empirical results and Section 6 concludes.

2 Clustering and productivity: underlying mechanisms

We consider two main mechanisms through which clustering could impact on the performance of firms: a competition mechanism and a spillover mechanism. First, consistent with the standard Cournot competition result, which suggests that the greater the number of firms in a market the lower the price, we expect that firms located in clusters with large numbers of firms in the same sector will face tougher competition and must operate more efficiently in order to survive. These firms will be forced to cut slack and use costs more efficiently and as a result should, ceteris paribus, appear more productive than firms located in markets with fewer local competitors. However, as is common in developing country settings, the adjustment process may be slow and competition in prices may first lead to an erosion of mark-ups before significant exit of firms is observed. Moreover, small-scale informal firms are unlikely to exit production even if they are loss-making as many are household enterprises with diversified income sources that may act as substitutes and serve both as sources of revenue and as risk coping mechanisms. As such, more intense competition in clusters may in fact lead to negative effects on measured firm performance.

Second, we consider whether firms experience spillover effects from other firms producing the same output that are located in close proximity. These are commonly referred to as localization economies.[5] The extent of such spillovers will depend on the characteristics of the cluster but also the characteristics of the firm. For example, we would expect to see positive external effects in the form of technology transfers for firms located within more productive clusters that are producing similar products. This could happen through the movement of labour between firms where employees take with them the technology and work practices of the firm they are leaving. Technological complementarities could also emerge where it is to the benefit of all firms in a cluster to adopt a new technology in which case technology diffusion is more likely. Fafchamps and Söderbom (2014) use the introduction of invoicing of clients as an example; such an innovation will only improve productivity and business practices if all firms selling to the same clients adopt the new technology. Spillovers can also take place through the copying or sharing of technologies diffused through local networks. Where firms' customers are based in the same locality as where they are producing, the sharing of technology is less likely and firms will have a greater incentive to protect whatever productivity advantage they might have. This type of diffusion, therefore, is more likely in large firms that serve markets outside of their immediate locality. It is also possible that some technologies are strategic substitutes whereby when one firm adopts a particular technology there is less of an incentive for others to do so in which case negative productivity spillovers are possible (Fafchamps and Söderbom 2014).

It is also the case that firms may have different absorptive capacities. This has been emphasized in the literature that explores the existence of spillovers from foreign direct investment (see Crespo and Fontoura (2007) and Marin and Bell (2006), for examples). It is possible that firms of different types may benefit differently from agglomeration and so in understanding the effects of clustering the characteristics of firms should be taken into account. For example, it may be that benefits are confined to large formal firms who have the scale and experience to benefit from knowledge or other productivity spillovers. It is also possible that large or formal firms are less flexible than small or informal firms and so are less likely to make the changes to production needed to benefit from such spillovers. Moreover, informal firms or small firms may compete more rigorously for the same local customers and so may go to greater lengths to prevent

[5] See, for example, Rosenthal and Strange (2004) and Overman and Venables (2005). In contrast, urbanization economies are benefits of locating in an area with more economic activity generally and are not related to the specific sector in which a firm operates. We control for urbanization economies in our analysis and focus on localization effects.

spillovers to competitors. However, as other literature has shown, informal firms in developing countries are often well-organized in business and other informal networks which make it more likely for spillovers from clustering to be realized (Overman and Venables 2005).

3 Empirical approach

Our aim in this paper is to identify the extent to which there are productivity gains associated with the clustering of firms using the case of Cambodia and to isolate, in particular, the competition and productivity spillover channels. We also aim to explore whether these effects vary by firm characteristics focusing in particular on registration/formality, manufacturing versus services sector firms, and the size of the firm. The approach we use is to estimate a firm level production function which allows us to relate the characteristics of the cluster the firm is located in with the productivity of the firm.

Empirically identifying the effects of clustering on firm performance is difficult due to the 'reflection problem' where it is difficult to separate out correlations in the productivity levels in clusters that are due to competition or spillover effects from correlated effects that are as a result of common shocks associated with other unobserved factors (Manski 1993). Separating out these effects is all the more challenging when exploiting cross-sectional variations in data.

We first construct a measure of the density of firms within clusters to control for the fact that firms are likely to locate in naturally advantageous areas such as urban centres or close to ports and other infrastructure. The inclusion of this variable will capture the extent to which firms are more productive when located in areas with a large amount of economic activity generally and will control for this in identifying competition and spillover effects. We also include a series of controls for the population density of the cluster and the quality of local infrastructure.

To capture competition effects we include the proportion of firms in the cluster that are in the same sector. This measure will establish whether being close to similar firms impacts on firm performance. A positive coefficient would indicate that there are mutual benefits from locating close to competitors, through for example, access to a common pool of labour, while a negative coefficient suggests that price competition effects dominate leading to lower measured performance of firms.

To capture productivity spillover effects we consider two further measures. First, we measure the average productivity of all other firms in the cluster, which, once the density of firms in the cluster is controlled for, will capture whether being in a high productivity cluster generally impacts on firm level performance. The inclusion of this variable will control to a certain extent for selection bias in that more productive firms are likely to locate close to other productive firms to benefit from the same types of agglomeration advantages such as access to suppliers or customers. Finally, to capture within-sector productivity spillovers we use the average productivity of all other firms in the cluster that are in the same sector. If firm performance is higher in clusters where there are more productive firms in the same sector this will provide evidence of spillover effects given that the density of the cluster, competition effects and selection effects are controlled for through the inclusion of the other measures.

In this model, it is still possible that a positive correlation between firm level productivity and cluster productivity is due to correlated effects. To address this problem we also include two controls for common shocks to firms in the cluster: (i) the change in the total size of the cluster (number of firms) between 2009 and 2011 (available years of data on clusters) and (ii) the proportion of firms in the cluster that come from the same sector. A change in the size of a

cluster or in the importance of a particular sector within a cluster is suggestive of a positive or negative shock common to all firms in that cluster thereby controlling for correlated effects that underpin the reflection problem. For each firm, we also compute cluster level productivity by excluding the information on the individual firm in question to minimize reverse causation due to the mechanical construction of the variables.

Our full empirical model is given by

$$\begin{aligned} ln\,out_{isj} &= \beta_0 + \beta_1 density_j + \beta_2 propfirm_{sj} + \beta_3 avprod_j + \beta_4 avprod_{sj} \\ &+ \beta_5 \Delta density_j + \beta_6 \Delta dpropfirm_{sj} + \boldsymbol{\alpha}\mathbf{X}_j + \boldsymbol{\delta}\mathbf{Z}_{isj} + \theta_s + \varphi_r + e_{isj} \end{aligned} \quad (1)$$

Where $ln\,out_{isj}$ is the output of firm *i* in sector *s* in cluster *j*; $density_j$ is the number of firms in cluster *j*; $propfirm_{sj}$ is the proportion of firms in sector *s* in cluster *j*; $avprod_j$ is the average productivity of firms in cluster *j*; $avprod_{sj}$ is the average productivity of firms in sector *s* in cluster *j*; $\mathbf{X}_j$ is a vector of cluster-specific control variables including population and measures of the quality of local infrastructure; $\mathbf{Z}_{isj}$ is a vector of firm-specific control variables including inputs and firm characteristics; θ_s are sector-specific fixed effects; φ_r are region-specific fixed effects; and e_{isj} is a statistical noise term. Δ indicates the change in a variable between 2009 and 2011.

The extent to which competition and productivity effects are identified empirically will also depend on how productivity is measured. In this paper output is based on the revenue of the firm. Revenue-based productivity measures capture both productivity differences and differences in mark-ups across firms. In competitive sectors the model given in Equation (1) will allow us to identify the effect of agglomeration on productivity given that firms will be operating with zero mark-ups. In non-competitive sectors where there is variability in mark-ups our model will pick up both productivity effects and the extent to which agglomeration places competitive pressure on markets and erodes such mark-ups. To address this issue as a robustness check on our results we consider the different effects in competitive and concentrated sectors to eliminate the possibility that any observed productivity effects are due to changes in mark-ups as opposed to real productivity improvements (Amiti and Konings 2007). Sector level concentration is measured using the standard Herfindahl-Hirschman index (HHI) as follows

$$HHI_s = \sum_{i=1}^{n} s_{is}^2 \quad (2)$$

where s_{is} is the revenue share of firm *i* in sector *s*. Our empirical model is adapted to take account of sector level concentration through the inclusion of interaction terms between the various clustering measures and the HHI.

4 Data

Our data come from the Cambodian nation-wide establishment listing gathered in 2009 (EL2009) and the Cambodian Economic Census conducted in 2011 (EC2011). The EL2009 covers all fixed establishments in Cambodia while the EC2011 covered both fixed and movable

establishments.[6] Both databases record the village in which the firm is located. Establishments classified in Sector A (agriculture, forestry, and fishing), Sector O (public administration and defence and compulsory social security), Sector T (activities of households as employers and producers of undifferentiated goods, and service-producing activities of households for their own use), and Sector U (activities of extra-territorial organizations and bodies) are excluded.

The EC2011 was conducted to provide fundamental statistics on the current status of business activities in Cambodia. Along with financial information on firms it includes information on the characteristics of the owners, the legal form of the firm, the nationality of the owner and manager, characteristics of employees, etc. The EL2009, however, only contains basic information on firms (type of enterprise, sector, and numbers employed) as its purpose was primarily to develop a census frame for the EC2011. A total of 376,761 establishments are covered by the EL2009 employing a total of 1,469,712 individuals. The EC2011 includes extensive information on 505,134 establishments employing a total of 1,676,263 individuals. The latter is the primary source used in this analysis.

Table 1 presents summary statistics on the key variables of interest based on EC2011. On average, there are 3.32 employees per establishment in 2011 highlighting the small average size of firms generally in Cambodia. In total 80 per cent of firms employ less than two people. Only 13,170 establishments employ ten or more with 12,383 of those having a labour force of between 10 and 99 people and only 787 firms with more than 100 employees. The majority of firms (98 per cent) are single unit firms and most (86 per cent) utilize premises that are less than $50m^2$ in area.

The majority of firms are service sector firms (85 per cent). There are 75,031 firms in the manufacturing sector in 2011 employing 539,134 people. The average number of persons engaged in manufacturing firms is 7.19 which is more than twice the average number of persons engaged in services (2.6) in 2011. Thirty-six per cent of all firm owners are male while 55 per cent of owners of manufacturing firms are male. The majority of female owned firms are small firms with less than 10 employees – 81 per cent of firms with 10 employees or more are owned by males.

Only 8 per cent of all firms in 2011 are registered to one of the ministries or agencies which approve the business operation of firms (5.6 per cent of manufacturing firms) highlighting that most of the firms operate in the informal sector of the economy. Moreover, 65 per cent of firms are categorized as home businesses (88 per cent of manufacturing firms) that are located in the residence of the owner. Fifteen per cent of firms are located in urban areas (12 per cent of manufacturing firms).

[6] The definition of an establishment by ISIC Rev4 is an economic unit that engages under a single ownership or control, that is, under a single legal entity, in one, or predominantly one kind of economic activity at a single physical location, for example, a mine, factory, or workshop. There are three kinds of establishments: fixed, movable, and mobile. Fixed means an establishment engaged in some form of economic activity in one fixed place; movable means that the establishment is fixed but that it is possible to move the location easily; while mobile establishments have no fixed place of business, for example street peddlers.

Table 1: Data description

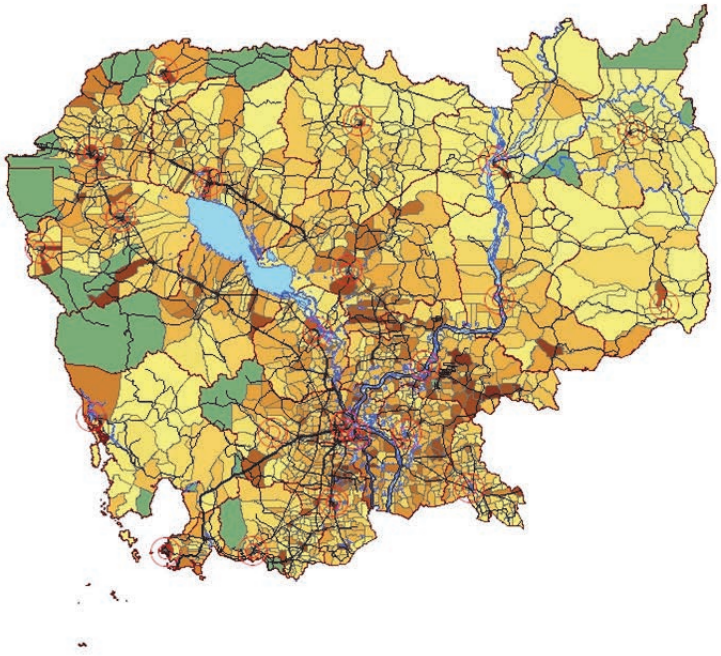
Characteristics of firms in the Enterprise Census 2011		
Number of firms	505,134	
Average numbers employed	3.32	
Proportion of firms employing <3	80%	
Proportion of firms employing 10-99	2%	
Proportion of firms employing 100+	0.2%	
Proportion of single unit firms	98%	
Proportion of premises <50m^2	86%	
Proportion male owners	36%	
Proportion formal	8%	
Proportion home business	88%	
Proportion located in urban area	15%	
Proportion manufacturing sector	15%	
Average numbers employed in manufacturing	7.19	
Proportion male owners in manufacturing	55%	
Proportion formal in manufacturing	5.6%	
Proportion home business in manufacturing	88%	
Proportion located in urban area in manufacturing	12%	
Characteristics of clusters in the Enterprise Listing 2009 and the Enterprise Census 2011		
	2009	2011
Number of firms in the village	206	308
Proportion of firms in the village in the same sector	0.247	0.217
Number of firms in the commune	618	967
Proportion of firms in the commune in the same sector	0.184	0.152

Source: authors' calculations using the Cambodian Economic Census 2011.

Table 1 also describes the nature of the clusters considered in our analysis for both 2011 and 2009. On average there are 308 firms in each village in 2011 compared with 206 firms in 2009. The average proportion of firms in the same ISIC4 sector in each village is 22 per cent in 2011 and 25 per cent in 2009. This suggests a high concentration of business activities within villages in both years. At the commune level, there are on average 967 firms in 2011 and 618 in 2009 (the average commune in Cambodia consists of 8.64 villages). On average 15 per cent of firms within commune clusters operate in the same ISIC4 sector in 2011 compared with 18 per cent in 2009. As outlined in Section 3, we use changes in the size and concentration of clusters to control for common shocks to firms within clusters in our empirical analysis.

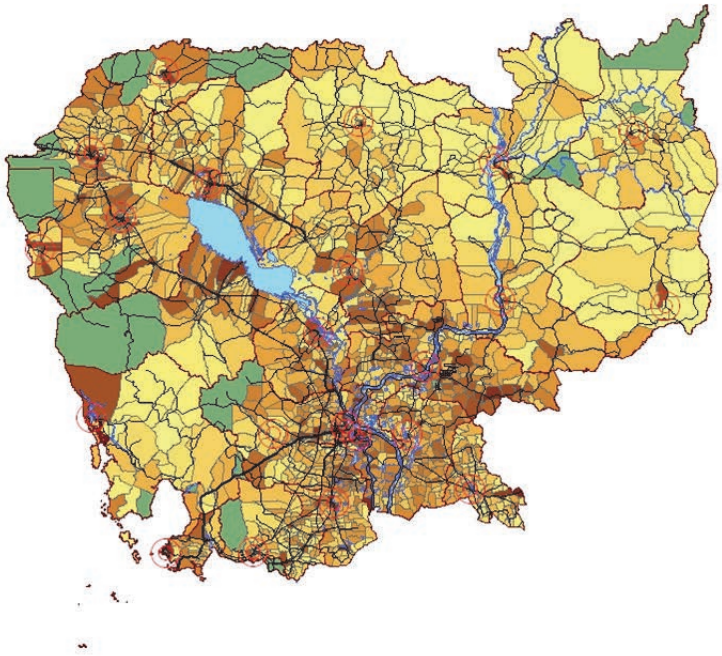
To describe the pattern of clustering in Cambodia we generate maps that illustrate the density of economic activity across communes in 2011 in terms of the number of firms (Figures 1a) and the numbers employed (Figure 1b).

Figure 1a: The location pattern of firms in Cambodia measured in terms of number of firms per district



Source: authors' calculations using the Cambodian Economic Census 2011.

Figure 1b: The location pattern of economic activity in Cambodia measured in terms of numbers employed per district



Source: authors' calculations using the Cambodian Economic Census 2011.

The highest density of firms is evident in provinces along Tonle Sap Lake, in Phnom Penh, and Kandal province (Phnom Penh surrounding province) and in the southern provinces where the population density is highest. There are very few firms located in the north east, northern, and south western regions. Comparing the number of firms with the numbers employed it is unsurprising that employment density closely matches the density as measured through the number of firms. One exception is in the north eastern region of Rattankiri and Mondolkiri provinces where the employment density is more concentrated. This suggests that establishments located in these regions are generally larger in size. The largest establishments, however, are mostly located in the urban centres of Phnom Penh, Kandal, Kompong Cham, Siem Reap, Sihanouk Ville, and Svay Rieng.

5 Empirical results

The variables included in our empirical model are described in detail in the Appendix. Output is measured as the log of annual sales. The main input variable is the log of the number of employees. Data on capital are not available and so we proxy for capital using the physical size of the premises where the firm operates (and its square) and a control variable for whether the firm is a single or multiple unit. We also control for the legal form of the firm, the ownership type and whether the firm is formally registered. We estimate the model by separately including both province and district level controls and also control for the sector that the firm operates in using 3-digit ISIC codes. We measure clustering at two levels of aggregation, the village and the commune. Standard errors are clustered at either the village or commune level depending on the level at which clustering is defined. Village and commune level measures of local population (number of households) and the quality of infrastructure are also included as controls. These variables are constructed using the 2008 Population Census. Two infrastructure quality variables are considered: (i) the proportion of households in the cluster (village or commune) with on-grid electrical power; and (ii) the proportion of households in the cluster with a piped water supply.[7] The results of the baseline specifications are presented in Table A2 of the Appendix.

In what follows, we focus on the relationship between clustering and productivity and on disentangling the various mechanisms through which clustering can impact on firm performance. We address four key questions: (i) are firms more or less productive where there is greater clustering of economic activities?; (ii) are different types of firms impacted differently by the clustering of economic activities?; (iii) are there productivity spillovers associated with clustering?; and (iv) Are different types of firms affected differently by productivity spillovers? We conclude the section by performing robustness checks on our key results.

5.1 Are firms more productive where there is greater clustering of economic activities?

To answer this question we include two measures of clustering in the regression models: the number of firms in the cluster and the proportion of firms in the cluster that are in the same sector. As discussed in Section 3, the former controls for the fact that firms generally are likely to cluster in naturally advantageous areas which will mean that they are likely to be more productive when located in areas with a large amount of economic activity generally. The inclusion of the population density of the cluster and the quality of infrastructure computed using the population census will also help to control for these factors. The proportion of firms in the cluster that are

[7] Out of the 14,584 villages in EC2011, 326 villages and 27 communes could not be matched to the 2008 Population Census. These villages and communes are excluded from the analysis reducing the total number of firms to 489,816. Their exclusion does not change our overall findings.

in the same sector is included to establish whether being located close to many competitors has an effect on firm performance and therefore captures potential competition effects. The results are presented in Table 2.

Table 2: Are firms more productive where there is more clustering of economic activity?

Dependent variable	lnsales				lnlabprod			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of firms in cluster	0.0001**	0.0001***	0.0001***	0.0001***	0.0001*	0.0001***	0.0001**	0.0001***
	(0.0001)	(0.00004)	(0.00003)	(0.00002)	(0.0001)	(0.00004)	(0.00003)	(0.00002)
Proportion of firms in same sector	-0.281***	-0.284***	-0.233***	-0.193***	-0.297***	-0.297***	-0.302***	-0.265***
	(0.038)	(0.034)	(0.069)	(0.060)	(0.039)	(0.035)	(0.070)	(0.061)
Regional fixed effects	Province	District	Province	District	Province	District	Province	District
Clustering	Village	Village	Commune	Commune	Village	Village	Commune	Commune
R-squared	0.379	0.397	0.374	0.391	0.572	0.583	0.570	0.581
n	489,816	489,816	497,022	497,022	498,552	498,552	505,943	505,943

Note: Each model includes controls for cluster population and infrastructure quality, firm characteristics and industry fixed effects. Levels of regional fixed effects and clustering are indicated in the table. Robust clustered standard errors presented in parenthesis. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.

Source: authors' calculations using the Cambodian Economic Census 2011.

We find that the greater the number of firms in a village or commune the higher the measured productivity of firms. The effect is similar when clustering is defined at both the village and commune level and is robust to the inclusion of both province and district fixed effects. The result is also similar when a simple measure of labour productivity is used as the left-hand side variable and suggests that firms will be more productive in locations where there is more economic activity. The inclusion of this measure controls for the possibility that clustering is associated with natural advantages such as infrastructure or a large population and allows us to identify whether there are additional effects associated with competition or spillovers within sectors. [8]

Turning to our sector-specific measure we find that the proportion of firms in the village/commune that are in the same sector has a negative effect on the productivity of firms. This suggests that the clustering of similar economic activities creates competitive pressures that lead to lower revenue levels for a given input level leading to an erosion of mark-ups or a decline in productivity. This provides evidence that clustering increases competition between firms located close together which in turn negatively affects measured firm performance.

5.2 Are different types of firms impacted differently by the clustering of economic activities?

To explore whether different types of firms are impacted differently by clustering we consider a range of interaction terms between firm characteristics and the clustering measures. These are

[8] The coefficients on the population density and infrastructure quality variables remain positive and well determined in each specification.

presented in Table 3. For ease of exposition we just present the results for the core regression model with district fixed effects. In all cases the results are confirmed when we consider labour productivity as the dependent variable and use province fixed effects.

Table 3: Which firms are more productive where there is clustering of economic activity?

Dependent variable lnsales	(1)	(2)	(3)	(4)	(5)	(6)
Number of firms in cluster	0.0001***	0.0001***	0.00005	0.0001***	0.0001***	0.0004
	(0.00004)	(0.00004)	(0.00005)	(0.00002)	(0.00002)	(0.00003)
Proportion of firms in same sector	-0.265***	-0.190***	-0.416**	-0.179***	0.014	-0.808**
	(0.035)	(0.038)	(0.209)	(0.061)	(0.067)	(0.339)
Registered	0.414***			0.398***		
	(0.024)			(0.035)		
Registered* number	0.000			0.00001		
	(0.000)			(0.00001)		
Registered* proportion	-0.681***			-1.138***		
	(0.144)			(0.364)		
Manufact[a]* number		-0.0002***			-0.00005**	
		(0.00005)			(0.00002)	
Manufact[a]* proportion		-0.415***			-0.692***	
		(0.078)			(0.150)	
Small			-0.716***			-0.762***
			(0.030)			(0.040)
Small* number			0.0001			0.00004
			(0.0001)			(0.00002)
Small* proportion			0.140			0.625*
			(0.210)			(0.338)
Regional fixed effects	District	District	District	District	District	District
Clustering	Village	Village	Village	Commune	Commune	Commune
R-squared	0.397	0.397	0.400	0.391	0.392	0.395
n	489,816	489,816	489,816	497,022	497,022	497,022

Notes: [a] The level effect of the indicator for manufacturing firms is absorbed by the ISIC3 digit fixed effects. Each model includes controls for cluster population and infrastructure quality, firm characteristics, and industry fixed effects. Levels of regional fixed effects and clustering are indicated in the table. Robust clustered standard errors are presented in parenthesis. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.
Source: authors' calculations using the Cambodian Economic Census 2011.

Given the large number of informal firms in Cambodia we first consider whether the impact of clustering is different for registered vs. unregistered firms. We include an indicator for whether the firm is registered to capture productivity differences between formal and informal firms and also an interaction term with each clustering measure (column 1). We find that registered firms are more productive than unregistered firms as is consistent with a growing empirical literature documenting the productivity gains of formalization (for example, Rand and Torm 2012). We also find that clustering in general does not impact on registered and unregistered firms in different ways as indicated by the insignificance of the interaction between the registered firm indicator and the density of firms in the cluster. We do find, however, that the negative effect of the proportion of firms in the cluster that are in the same sector is much more pronounced for

registered firms. This suggests that competition arising from the clustering of similar firms has a strong negative effect on firm performance for formal enterprises in particular.

One possible explanation for the fact that formal firms appear to be more negatively impacted by competition is that they may face a higher and less flexible cost base due to taxes and other charges that they must pay once they register. In the face of increased competitive pressures they may find it more difficult to cut costs and so suffer greater losses than unregistered firms that have greater flexibility. This explanation is consistent with the nature of the formalization process in Cambodia. In the Cambodian case the formalization process is tedious and cumbersome taking on average 104 days (compared with 37 days for the East Asia and Pacific region as a whole) and costing on average 151 per cent of the annual per capita income (compared with 29.8 per cent on average for the East Asia and Pacific region) (World Bank 2009). Moreover, results from the 2007 Investment Climate Survey suggest that formal enterprises face more stringent legal and institutional constraints while informal enterprises are less likely to report corruption as a constraint (World Bank 2009). Our findings for Cambodia are also consistent with McKenzie and Sakho (2007) who find that business registration reduces the profits of some firms. Our results suggest that one of the underlying mechanisms may be that formal firms are more inflexible in adjusting the cost base in the face of increased competition compared with informal firms.

We also distinguish between manufacturing and service sector firms (column 3). We find that while all firms are more productive in clusters that are more densely populated with other firms, manufacturing firms are impacted to less of an extent. This suggests that service sector firms, in particular, do better where there is a concentration of economic activity. This result is likely due to the fact that most service firms need to locate close to their customers and there are naturally more potential customers in areas of high economic activity. We also find that competition effects are more pronounced for firms in the manufacturing sector. While service sector firms also perform worse in clusters with more competitors, for manufacturing firms the effect is about twice the magnitude. This is contrary to our expectation that service firms are more vulnerable in settings where local competition is more intense as compared with manufacturing firms given that for many service firms their customers must be located at the point of sale. Our results suggest that while there are negative competition effects, these effects are even more pronounced for manufacturing firms. One possible explanation is that there are other factors impacting on the ability of manufacturing firms to diversify their customer base by selling their products outside of their local area. These factors could be, for example, high transport costs, a lack of market information or the small scale of operations (see, for example, Siba et al. 2012).

As discussed in section 4, the majority of firms in Cambodia are small with less than 10 employees. To uncover whether they are impacted differently than larger firms by clustering and competition effects we include an interaction term between an indicator for small firms and the clustering measures (column 5). We find no strong evidence to suggest that firms of different sizes are impacted in different ways by competition effects associated with clustering.

5.3 Are there productivity spillovers associated with clustering?

As discussed in Section 3, identifying empirically the productivity spillover effects associated with clustering is difficult with cross-sectional data. To tighten our identification of these effects we include a number of control variables. The inclusion of the average productivity of all other firms in the cluster will capture whether being in a high productivity cluster generally impacts on firm level performance and will control to a certain extent for the selection of more productive firms into more productive locations or in locations with greater agglomeration advantages such as access to suppliers or customers. We also include controls for common shocks to firms in the

cluster by controlling for the change in the total size of the cluster (number of firms) between 2009 and 2011 and the proportion of firms in the cluster that come from the same sector. Once all of these factors are included, we can identify productivity spillover effects through the average productivity of firms in the cluster in the same sector. It should be noted that the productivity of the individual firm in question is excluded in computing the cluster level productivity measures. We estimate the model for clustering defined at both the district and the commune level. The results are presented in Table 4.

Table 4: Are firms more productive in more productive clusters?

Dependent variable lnsales	(1)	(2)	(3)	(4)
Average prod. of firms in cluster	0.001	0.001	0.046***	0.039***
	(0.001)	(0.001)	(0.011)	(0.009)
Average prod. of firms in same sector	0.001	0.001	0.001	0.001**
	(0.001)	(0.001)	(0.001)	(0.001)
Number of firms in cluster	0.00001	0.00004	-0.00002	0.000
	(0.0001)	(0.00004)	(0.00003)	(0.000)
Proportion of firms in same sector	-0.316***	-0.316***	-0.185***	-0.136**
	(0.046)	(0.040)	(0.070)	(0.060)
Change in number of firms in cluster	0.0005***	0.0004***	0.0003***	0.0003***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Change in proportion of firms in same sector	0.001	0.005	-0.202***	-0.197***
	(0.041)	(0.035)	(0.036)	(0.032)
Regional fixed effects	Province	District	Province	District
Clustering	Village	Village	Commune	Commune
R-squared	0.382	0.398	0.380	0.395
n	489,238	489,238	497,022	497,022
Marginal effects				
1% increase in number of firms in same sector	-0.001	-0.012	0.452	0.403
1% increase in number of firms in diff sector	0.002	-0.009	0.456	0.406
Net agglomeration effect	-0.003	-0.003	-0.004	-0.003

Notes: Each model includes controls for cluster population and infrastructure quality, firm characteristics, and industry fixed effects. Levels of regional fixed effects and clustering are indicated in the table. Robust clustered standard errors presented in parenthesis. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.

Source: authors' calculations using the Cambodian Economic Census 2011 and Cambodian Establishment Listing 2009.

When clustering is defined at the village level we find very little evidence of productivity spillovers from clustering. Moreover, the large negative competition effect identified in Table 2 persists even when changes in the size and composition of clusters are controlled for. This suggests that the competition induced by clustering is a detrimental factor in determining firm performance.

To quantify the impact we consider the marginal effect of a 1 per cent increase in the number of firms in a cluster assuming that the new firms are at the average productivity level of incumbents.

We consider the differential impact on average productivity when the new firms are part of the same sector as compared with when they are not. The difference in the two marginal effects will allow us to isolate the effect that can be attributed to the agglomeration of firms in the same sector. The effects are presented at the bottom of Table 4. The overall net effect is a small negative impact of between 0.3 and 0.4 per cent. In other words, the combination of competition effects and spillover effects due to the agglomeration of firms in the same sector have only a very small overall impact on the productivity of firms when the other effects are netted out. This is similar to findings by Siba et al. (2012) who also find that the competition effects of clustering outweigh any productivity spillovers.

When clustering is defined at the commune level we find some evidence of positive productivity spillovers once we control for district fixed effects (column 4). The negative effect of competition from clustering (captured by the proportion of firms in the same sector), however, is still negative and well determined. This suggests that in more broadly defined geographical areas, where there is less competition between firms, there is greater potential for productivity spillovers. However, on net, the overall marginal effect is negative.

5.4 Are different firms affected differently by productivity spillovers?

As a final step we disentangle which firms are most affected by being in productive clusters by disaggregating the analysis by the characteristics of the firm. Table 5a presents the results for village level clustering while the results for commune level clustering are presented in Table 5b.

Table 5a: Which firms are more productive in more productive clusters (village level clustering)?

Dependent variable	(1)	(2)	(3)	(4)	(5)
lnsales	Registered	Unregist.	Manufact.	Services	Small
Average prod. of firms in cluster	0.009***	0.0001	0.019***	0.0006	0.001
	(0.002)	(0.001)	(0.003)	(0.001)	(0.001)
Average prod. of firms in same sector	-0.001	0.002**	0.041***	0.001	0.001
	(0.002)	(0.001)	(0.003)	(0.001)	(0.001)
Number of firms in cluster	0.0002***	0.00005	-0.0001*	0.00004	0.00005
	(0.0001)	(0.00004)	(0.00005)	(0.00004)	(0.00004)
Proportion of firms in same sector	-0.988***	-0.278***	-0.865***	-0.150***	-0.296***
	(0.210)	(0.040)	(0.084)	(0.042)	(0.040)
Change in number of firms in cluster	0.0001	0.0004***	-0.0002	0.0004***	0.0004***
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Change in proportion of firms in same sector	-0.077	-0.007	0.286***	-0.097**	-0.003
	(0.184)	(0.035)	(0.065)	(0.041)	(0.035)
R-squared	0.595	0.366	0.482	0.336	0.363
n	36,101	453,137	67,197	417,104	480,009
Marginal effects					
1% increase in number of firms in same sector	0.011	-0.005	0.121	0.027	-0.008
1% increase in number of firms in diff sector	0.022	-0.002	0.126	0.030	-0.005
Net agglomeration effect	-0.010	-0.003	-0.006	-0.002	-0.003

Notes: Each model includes controls for cluster population and infrastructure quality, firm characteristics, and industry fixed effects. Robust standard errors clustered at the village level are presented in parenthesis. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.

Source: authors' calculations using the Cambodian Economic Census 2011 and Cambodian Establishment Listing 2009.

Table 5b: Which firms are more productive in more productive clusters (commune level clustering)?

Dependent variable:	(1)	(2)	(3)	(4)	(5)
lnsales	Registered	Unregist.	Manufact.	Services	Small
Average prod. of firms in cluster	0.005	0.046***	0.097***	0.031***	0.040***
	(0.011)	(0.010)	(0.016)	(0.008)	(0.009)
Average prod. of firms in same sector	0.002**	0.002*	0.014	0.001**	0.001**
	(0.001)	(0.001)	(0.011)	(0.001)	(0.001)
Number of firms in cluster	0.0001*	0.000	-0.00006	0.000	0.000
	(0.00004)	(0.000)	(0.00004)	(0.000)	(0.000)
Proportion of firms in same sector	-1.137**	-0.105*	-0.338***	0.090	-0.114*
	(0.526)	(0.061)	(0.131)	(0.066)	(0.060)
Change in number of firms in cluster	0.0001	0.0003***	0.0001	0.0003***	0.0003***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Change in proportion of firms in same sector	-0.253	-0.203***	-0.043	-0.260***	-0.200***
	(0.160)	(0.033)	(0.063)	(0.038)	(0.032)
R-squared	0.594	0.362	0.477	0.332	0.359
n	36,557	460,465	68,205	423,837	487,614
Marginal effects					
1% increase in number of firms in same sector	0.052	0.471	0.808	0.381	0.413
1% increase in number of firms in diff sector	0.065	0.474	0.812	0.383	0.416
Net agglomeration effect	-0.014	-0.003	-0.004	-0.002	-0.003

Notes: Each model includes controls for cluster population and infrastructure quality, firm characteristics, and industry fixed effects. Robust standard errors clustered at the commune level are presented in parenthesis. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.

Source: authors' calculations using the Cambodian Economic Census 2011 and Cambodian Establishment Listing 2009.

Columns 1 and 2 in Tables 5a and 5b present the results for the model estimated separately for registered and unregistered firms. For clustering at the village level, we do not find any evidence to suggest that registered firms experience productivity spillovers from other firms located in the same village. There is some evidence to suggest that there are productivity spillovers between registered firms within communes. This suggests that knowledge spillovers are less likely between firms that are located very close together but exist between firms that have some distance between them. As was found in Table 3, registered firms are significantly negatively affected by the clustering of competitors through the competition channel, both within village and commune level clusters. The negative effect of competition outweighs any gains in productivity due to spillovers. The marginal effect is larger than for the overall sample at -1 per cent, and is even larger when clustering is defined at the commune level. Overall, for formal registered enterprises agglomeration has a negative effect on productivity.

For informal enterprises (column 2) we find a positive and significant relationship between the productivity of other firms in the same sector in the cluster and the productivity of the firm suggesting that there are productivity spillovers for informal firms, both within villages and communes. However, the negative competition effect persists. The overall marginal effect attributed to agglomeration is negative but of a much lower magnitude than for registered firms due to both competition effects that are of a smaller magnitude and due to the role played by productivity spillovers. The existence of positive productivity spillovers for informal firms

suggests that there are gains to informal firms from locating close together. As proposed by Overman and Venables (2005) these could be due to spillovers from a concentration of labour or knowledge that transfers between firms, both formal and informal. While on net the impact is negative, our results suggest that the informal sector in Cambodia is dynamic and that net gains from agglomeration are possible.

Next we isolate manufacturing versus service sector firms (columns 3 and 4). For manufacturing firms we find positive productivity spillovers when clustering is defined at the village level (Table 5a). The magnitude of this effect, however, is not large enough to outweigh the negative competition effects associated with being in a densely populated cluster of same-sector firms. The overall marginal effect is negative and larger than that found when all firms are considered. When the definition of a cluster is broadened to encompass the commune, the productivity effect disappears (Table 5b), but so too do some of the negative competition effects and the overall marginal effect is more in line with that found for the whole sample at -0.4 per cent. This suggests that the positive productivity spillovers that exist between manufacturing firms are more likely to be diffused through near neighbours than competitors in localities that are further away, but competition effects inhibit firms from realizing the full benefits of such spillovers.

For service sector firms, positive productivity spillovers are observed when clustering is defined at the commune level but not at the village level. This suggests that service sector firms benefit more from having some space between them and their competitors. This is not surprising given that service sector firms in particular are more likely to be in direct competition for customers with other service sector firms in the same village given that services are consumed at the point of sale. As such they are less likely to share technologies or learning with their very local competitors but are more likely to do so with similar firms in the same locality but not in direct competition with them. The overall effect is negative whether clustering is defined at the village or commune level but is smaller than for the whole sample.

Finally, considering the size of the firm, we find that for small firms productivity spillovers are evident within communes but not between firms located in the same village (column 5). It is likely that, as for service sector firms, small firms are more likely to compete very locally and so are less likely to learn from their local competitors but will learn from productive, non-competitor, firms doing similar activities in the nearby area.

5.5 Robustness checks

Our results provide evidence of a link between the clustering of economic activity in Cambodia and the productivity of firms through both competition and productivity spillover channels. On net, we find that the overall effect is negative but of a small magnitude. We also find that the nature of the relationship is different for different types of firms. We perform two robustness checks on these results. First, we add an additional check that our results are not driven by self-selection in the location choice decision, and second, we consider whether the effects hold in competitive sectors by including interaction terms with a sector level concentration measure.

Even though we control for common shocks to clusters through including variables that capture the change in the clustering of activity between 2009 and 2011, control for selection through the inclusion of general measures of the size and quality of clusters, and control for population density and infrastructure quality in the cluster, it may still be that more productive firms are more likely to locate in villages or communes that have the potential to yield productivity improvements. If this is the case then there may be an endogenous self-selection process at work that confounds our results. To address this issue we perform a robustness check where we limit our analysis to older firms, i.e. firms that were in existence in 2009. In this way we omit firms

that have made their location choice in a recent time period and so would have been more likely to self-select into that location on the basis of the current productivity levels of other firms in that location. The results are presented in Table 6.

All of our results remain robust to the estimation of the models using the reduced sample and the overall marginal effects are almost identical. This suggests that the effects we observe are unlikely to be driven by the self-selection of more productive firms into productive locations.

Our second robustness check attempts to disentangle real productivity spillovers from impacts on mark-ups by interacting each measure of clustering with the sector level concentration level computed using the HHI given in Equation (2). The interaction terms capture the impact of clustering on firms in more concentrated sectors and so the coefficients on the level effect are the parameter of interest. They measure the impact on the productivity of firms in competitive sectors where mark-ups are low and thereby any positive impacts can be attributed to productivity gains. The results are presented in Table 7.

The marginal effects presented relate to those for firms in competitive sectors (i.e. using the coefficients on the level effects). All of our core results hold and the magnitudes of the marginal effects are almost identical. This suggests that our findings in relation to the productivity impacts of clustering are due to impacts on real productivity as opposed to just mark-ups.

Table 6: Robustness check excluding new entrants since 2009

Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
lnsales			Reg.	Unreg.	Man.	Serv.	Small	Reg.	Unreg.	Man.	Serv.	Small
Average prod. of firms in	0.001	0.041***	0.009***	-0.0002	0.021***	0.001	0.002	0.009	0.049***	0.203***	0.032***	0.044***
cluster	(0.001)	(0.010)	(0.003)	(0.001)	(0.003)	(0.001)	(0.001)	(0.010)	(0.012)	(0.018)	(0.008)	(0.010)
Average prod. of firms in same	0.001	0.001	-0.001	0.002**	0.042***	0.001	0.0004	0.002**	0.001	0.012	0.001	0.001
sector	(0.001)	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.010)	(0.001)	(0.001)
Number of firms in cluster	0.00004	0.000	0.0001**	0.00005	-0.0001**	0.00004	0.00004	0.0001*	0.000	-0.0001*	0.000	0.000
	(0.00004)	(0.000)	(0.0001)	(0.00004)	(0.00005)	(0.00004)	(0.00004)	(0.00004)	(0.000)	(0.00004)	(0.000)	(0.000)
Proportion of firms in same	-0.368***	-0.185***	-0.962***	-0.329***	-0.862***	-0.195***	-0.347***	-1.167**	-0.152**	-0.302**	0.047	-0.164**
sector	(0.043)	(0.064)	(0.219)	(0.043)	(0.087)	(0.043)	(0.043)	(0.553)	(0.064)	(0.133)	(0.070)	(0.064)
Change in number of firms in	0.0004***	0.0003***	0.0002	0.0004***	-0.00003	0.0004***	0.0004***	0.00004	0.0003***	0.0001	0.0003***	0.0003***
cluster	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Change in proportion of firms	0.057	-0.159***	-0.058	0.045	0.324***	-0.052	0.048	-0.214	-0.164***	-0.017	-0.229***	-0.163***
in same sector	(0.038)	(0.036)	(0.185)	(0.038)	(0.067)	(0.043)	(0.038)	(0.151)	(0.036)	(0.066)	(0.043)	(0.035)
Clustering	Village	Commune	Village	Village	Village	Village	Village	Commune	Commune	Commune	Commune	Commune
R-squared	0.410	0.406	0.602	0.378	0.498	0.334	0.372	0.602	0.374	0.493	0.330	0.368
n	369,184	374,148	31,772	337,412	54,161	311,125	361,143	32,122	342,026	54,845	315,376	365,973
Marginal effects												
1% increase in number of firms in same sector	-0.023	0.409	0.019	-0.020	0.135	0.024	-0.015	0.067	0.476	1.617	0.382	0.436
1% increase in number of firms in diff sector	-0.020	0.412	0.029	-0.017	0.140	0.026	-0.012	0.081	0.479	1.620	0.384	0.439
Net agglomeration effect	-0.003	-0.003	-0.010	-0.003	-0.005	-0.002	-0.003	-0.014	-0.003	-0.003	-0.002	-0.003

Notes: Each model includes controls for cluster population and infrastructure quality, firm characteristics, industry, and district fixed effects. Robust clustered standard errors are presented in parenthesis. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.

Source: authors' calculations using the Cambodian Economic Census 2011 and Cambodian Establishment Listing 2009.

Table 7: Controlling for sector level concentration, results for competitive sectors presented.

Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
lnsales			Reg.	Unreg.	Man.	Serv.	Small	Reg.	Unreg .	Man.	Serv.	Small
Average prod. of firms in	0.0005	0.038***	0.006***	-0.0005	0.020***	-0.00002	0.001	0.0002	0.045***	0.104***	0.030***	0.039***
cluster	(0.001)	(0.009)	(0.002)	(0.001)	(0.003)	(0.001)	(0.001)	(0.011)	(0.011)	(0.018)	(0.008)	(0.009)
Average prod. of firms in	0.001	0.001	-0.001	0.002***	0.047***	0.001	0.001	0.002**	0.002*	0.017	0.001*	0.001
same sector	(0.001)	(0.001)	(0.002)	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)	(0.014)	(0.001)	(0.001)
Number of firms in cluster	0.00004	-0.00001	0.0002*	0.00004	-	0.00004	0.00004	0.0001	0.000	-0.0001***	0.000	0.000
	(0.00004)	(0.00003)	*	(0.00004)	0.0002***	(0.00004)	(0.00004)	(0.00004)	(0.000)	(0.00004)	(0.000)	(0.000)
			(0.0001)		(0.00005)							
Proportion of firms in	-0.314***	-0.142**	-	-0.283***	-1.003***	-0.142***	-0.301***	-0.825*	-0.119*	-0.642***	0.093	-0.130**
same sector	(0.040)	(0.066)	0.904***	(0.041)	(0.093)	(0.042)	(0.041)	(0.505)	(0.067)	(0.183)	(0.066)	(0.066)
			(0.215)									
Change in number of	0.0004***	0.0003***	0.0002	0.0004***	0.0001	0.0004**	0.0004***	0.00005	0.0003***	0.0002**	0.0003***	0.0003***
firms in cluster	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Change in proportion of	0.026	-0.179***	-0.070	0.011	0.341***	-0.076*	-0.017	-0.191	-0.192***	0.006	-0.233***	-0.185***
firms in same sector	(0.036)	(0.034)	(0.187)	(0.036)	(0.068)	(0.041)	(0.036)	(0.163)	(0.034)	(0.068)	(0.038)	(0.033)
HHI*Av prod. of firms in	0.027**	0.064***	0.213***	0.025	-0.015	0.062**	0.026*	0.312**	0.049*	-0.050	0.162*	0.050**
cluster	(0.014)	(0.025)	(0.035)	(0.015)	(0.021)	(0.032)	(0.015)	(0.101)	(0.026)	(0.061)	(0.087)	(0.024)
HHI*Av prod. of firms in	0.008	0.006	0.051**	0.010	-0.085***	0.056*	0.008	0.020***	0.002	-0.024	0.001	0.008
same sector	(0.012)	(0.009)	(0.022)	(0.013)	(0.021)	(0.030)	(0.012)	(0.005)	(0.014)	(0.032)	(0.014)	(0.010)
HHI*Number of firms in	0.0004	0.0001	-0.0001	0.0004	0.004***	-0.0004	0.0004	-0.00005	0.0001	0.0008***	0.0001	0.0001
cluster	(0.0002)	(0.0001)	(0.0003)	(0.0002)	(0.001)	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0003)	(0.0002)	(0.0001)
HHI*Prop of firms in	0.183	0.628	-5.055*	0.579	2.970***	0.156	0.629	-49.901***	0.937	3.904***	1.506	1.116
same sector	(0.468)	(0.794)	(2.828)	(0.460)	(0.601)	(2.520)	(0.482)	(8.364)	(0.800)	(1.196)	(3.830)	(0.795)
HHI*Ch in number of	-0.001	-0.004	-0.001	-0.001	-0.005*	-0.00004	-0.001	-0.0001	-0.0004	-0.0008	-0.001*	-0.0004
firms in cluster	(0.0005)	(0.0003)	(0.001)	(0.0006)	(0.003)	(0.0006)	(0.001)	(0.0004)	(0.0003)	(0.00006)	(0.0004)	(0.0003)
HHI*Ch in prop of firms in	-1.283***	-1.014**	2.023	-1.267***	-1.919***	-6.310***	-1.387***	-0.883	-0.698*	-1.232**	-6.300***	-0.942**
same sector	(0.398)	(0.436)	(2.861)	(0.382)	(0.521)	(2.260)	(0.395)	(2.004)	(0.426)	(0.570)	(1.734)	(0.422)

Clustering	Village	Commune	Village	Village	Village	Village	Village	Commune	Commune	Commune	Commune	Commune
R-squared	0.398	0.395	0.597	0.367	0.484	0.336	0.363	0.596	0.362	0.479	0.332	0.359
n	489,238	497,022	36,101	453,137	67,197	417,104	480,009	36,557	460,465	68,205	423,837	487,614
Marginal Effects												
1% increase in number of firms in same sector	-0.016	0.393	-0.007	-0.012	0.139	0.023	-0.009	0.010	0.460	0.847	0.373	0.403
1% increase in number of firms in diff sector	-0.013	0.397	0.003	-0.009	0.145	0.026	-0.006	0.020	0.463	0.853	0.375	0.406
Net agglomeration effect	-0.003	-0.003	-0.010	-0.003	-0.006	-0.002	-0.003	-0.010	-0.003	-0.006	-0.001	-0.003

Notes: Each model includes controls for cluster population and infrastructure quality, firm characteristics and industry, and district fixed effects. Robust clustered standard errors are presented in parenthesis. The level effect of HHI is included in each model. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.

Source: authors' calculations using the Cambodian Economic Census 2011 and Cambodian Establishment Listing 2009.

6 Conclusion

This paper explored the pattern of clustering of firms in Cambodia and investigated the extent to which the clustering of economic activity leads to productivity improvements for firms. We explored two competing channels through which firms could be affected by locating near their main competitors: (i) the competition channel whereby increased competition within clusters has a negative effect on firm performance; and (ii) the spillover channel where firms within clusters learn from each other and experience productivity gains due to their physical proximity.

We used Enterprise Census data for Cambodia on over 500,000 enterprises in both manufacturing and services sectors and in the formal and informal economy. Overall, our analysis suggests that once the density of economic activity is controlled for, there are negative competition effects associated with clustering. The firms facing the greatest competitive pressures are formal, registered enterprises and manufacturing firms. Surprisingly we find that competition effects are of a lower magnitude for service sector and informal firms. We also find some evidence of productivity spillovers particularly for informal firms and firms in the manufacturing sector. While some care should be taken in inferring causality in these relationships we include a number of important controls to tighten our identification strategy including controls for the quality of local infrastructure, selection bias, correlated effects and differential effect in competitive, compared with concentrated sectors.

Overall, our results suggest that while there are observed benefits to firm performance from the clustering of economic activity in Cambodia associated with productivity spillovers, these benefits do not outweigh the negative impact of competitive pressures resulting from similar firms locating in close proximity to each other. This is consistent with Siba et al. (2012) who also find evidence for negative competition effects and positive productivity spillovers associated with the clustering of firms in Ethiopia. Of particular note in the Cambodian case is that we find significant positive productivity spillovers associated with the clustering of informal firms. To our knowledge there is no other quantitative evidence available on the potential for agglomeration economies for informal firms in developing country contexts.

Even though the negative competition effect is likely to outweigh any productivity advantages of clustering, our findings suggest the existence of a dynamic informal sector in Cambodia capable of experiencing agglomeration economies. It also suggests that there are inflexibilities associated with formality that make it more difficult for firms to adapt and compete when exposed to competitive pressures. Similarly, our findings for manufacturing firms suggest that they face constraints in accessing customers in other localities which would ease the competitive pressures from other firms within the cluster. Both sets of constraints inhibit the ability of formal and manufacturing firms to benefit from the typical agglomeration economies often proposed in the industrial policies of developing nations. Given that industrial policy tends to be primarily directed towards these firms further investigation into the nature and extent of these constraints is warranted. While this paper does not consider the role of government formulated clusters such as export processing zones and industrial parks, on the basis of the 'naturally' formed clusters that we identify in our data we can conclude that the effectiveness of a policy that creates incentives for similar firms to locate near each other will depend on the removal of constraints that impact on the cost base of firms, particularly formal and manufacturing firms, that typically form the basis of a healthy industrial sector.

Appendix Table A1: Variable descriptions and descriptive statistics

Variable name	Description	Mean	Std. Dev.
Dependent variables			
lnsales	Log of annual sales	8.516	1.285
lnlabprod	Log of labour productivity (sales/numbers employed)	7.820	1.619
Independent variables			
lnlabour	Log of total numbers employed	0.574	0.683
register	Dummy = 1 if firm is registered with a ministry or agency	0.084	0.278
owner_foreign	Dummy = 1 if firm is owned by a foreign national	0.011	0.105
owner_male	Dummy =1 if firm is owned by a male	0.357	0.479
urban	Dummy = 1 if firm is in urban area	0.150	0.357
foreign	FDI firm	0.0002	0.013
state	State owned firm	0.024	0.153
Business type			
kind_1	Street business	0.082	0.274
kind_2	Home business	0.645	0.478
kind_3	Apartment building	0.027	0.161
kind_4	Traditional market	0.177	0.382
kind_5	Modern shopping centre	0.001	0.039
kind_6	One exclusive block/building	0.053	0.225
kind_7	Other	0.014	0.119
area	Total area of business in square metres	11.33	16.52
single	Dummy =1 if firm is one single unit	0.982	0.133
Cluster controls [a]			
Nr_vill	Number of households in the village	487	687
Pr_ongrid_vill	Proportion of households in the village with on-grid electricity	0.397	0.410
Pr_pipedwater_vill	Proportion of households in the village with piped water	0.262	0.360
Nr_comm	Number of households in the commune	3,000	2,868
Pr_ongrid_comm	Proportion of households in the commune with on-grid electricity	0.362	0.377
Pr_pipedwater_comm	Proportion of households in the commune with piped water	0.238	0.331
Cluster measures			
Nr_firm_vill	Number of firms in the village	308	552
Prop_firm_vill_sec	Proportion of firms in the village in the same sector	0.217	0.231
Nr_firm_comm	Number of firms in the commune	967	1,165
Prop_firm_comm_sec	Proportion of firms in the commune in the same sector	0.152	0.176
Lnlabprod_vill	Average labour productivity of firms in the village [b]	9.32	11.95
Lnlabprod_vill_sec	Average labour productivity of firms in the village in the same sector [b]	7.84	14.33
Lnlabprod_comm	Average labour productivity of firms in the village [b]	8.27	1.77
Lnlabprod_comm_sec	Average labour productivity of firms in the village in the same sector [b]	9.36	15.29

Notes: [a]Calculated using the 2008 Population Census; [b]Computed excluding labour productivity of firm *i*.
Source: authors' calculations using the Cambodian Economic Census 2011 and Cambodian Establishment Listing 2009.

Appendix Table A2: Baseline productivity regressions

Dependent variable	Lnsales				Lnlabprod			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
lnlabour	0.583***	0.586***	0.586***	0.590***				
	(0.007)	(0.007)	(0.010)	(0.009)				
register	0.350***	0.341***	0.360***	0.352***	0.160***	0.154***	0.170***	0.164***
	(0.022)	(0.021)	(0.025)	(0.024)	(0.020)	(0.020)	(0.024)	(0.023)
owner_foreign	0.156***	0.154***	0.165***	0.176***	0.023	0.024	0.032	0.043
	(0.040)	(0.036)	(0.046)	(0.040)	(0.034)	(0.032)	(0.038)	(0.033)
owner_male	0.001	-0.004	0.001	-0.002	-0.056***	-0.055***	-0.057***	-0.056***
	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)	(0.008)	(0.007)
urban	0.011	-0.006	0.001	0.001	0.019	-0.014	0.010	-0.007
	(0.026)	(0.030)	(0.037)	(0.038)	(0.027)	(0.029)	(0.038)	(0.039)
foreign	1.264***	1.181**	1.198***	1.078**	-0.144	-0.174	-0.154	-0.196
	(0.454)	(0.500)	(0.451)	(0.499)	(0.380)	(0.368)	(0.299)	(0.276)
state	-0.489***	-0.478***	-0.526***	-0.516***	-0.649***	-0.641***	-0.672***	-0.665***
	(0.098)	(0.090)	(0.103)	(0.098)	(0.075)	(0.070)	(0.078)	(0.076)
kind_2	-0.086***	-0.090***	-0.103***	-0.107***	-0.139***	-0.141***	-0.154***	-0.155***
	(0.016)	(0.016)	(0.017)	(0.016)	(0.017)	(0.016)	(0.017)	(0.017)
kind_3	-0.013	-0.005	-0.021	-0.012	-0.101***	-0.090***	-0.107***	-0.095***
	(0.029)	(0.029)	(0.034)	(0.034)	(0.029)	(0.030)	(0.033)	(0.033)
kind_4	0.360***	0.358***	0.393***	0.381***	0.350***	0.338***	0.380***	0.368***
	(0.032)	(0.027)	(0.031)	(0.027)	(0.033)	(0.028)	(0.032)	(0.028)
kind_5	0.460***	0.447***	0.471***	0.470**	0.112	0.112	0.127	0.135
	(0.091)	(0.146)	(0.118)	(0.184)	(0.102)	(0.165)	(0.109)	(0.181)
kind_6	-0.173***	-0.182***	-0.192***	-0.204***	-0.358***	-0.362***	-0.380***	-0.385***
	(0.054)	(0.052)	(0.059)	(0.057)	(0.044)	(0.042)	(0.048)	(0.046)
kind_7	-0.114***	-0.125***	-0.134***	-0.144***	-0.122***	-0.129***	-0.141***	-0.148***
	(0.034)	(0.032)	(0.035)	(0.034)	(0.033)	(0.032)	(0.034)	(0.034)
area	0.128***	0.132***	0.131***	0.134***	0.082***	0.084***	0.086***	0.087***
	(0.011)	(0.011)	(0.013)	(0.013)	(0.011)	(0.010)	(0.013)	(0.012)
area_squared	0.0004	0.0002	0.0002	0.0001	-0.0006	-0.0005	-0.001	-0.001
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
single	-1.163***	-1.148***	-1.175***	-1.160***	7.953***	7.968***	7.953***	7.968***
	(0.114)	(0.112)	(0.118)	(0.112)	(0.060)	(0.059)	(0.069)	(0.068)
Nr_vill	0.00005**	0.0001***			0.0005***	0.0001***		
	(0.00002)	(0.00002)			(0.0002)	(0.0002)		
Pr_ongrid_vill	0.407***	0.323***			0.370***	0.291***		
	(0.032)	(0.031)			(0.032)	(0.031)		
Pr_pipedwater_vill	0.193***	0.272***			0.165***	0.255***		
	(0.051)	(0.051)			(0.051)	(0.052)		
Nr_comm			0.000	0.0001**			0.000	0.00001*
			(0.000)	(0.00001)			(0.000)	(0.000)
Pr_ongrid_comm			0.502***	0.391***			0.463***	0.358***
			(0.061)	(0.061)			(0.062)	(0.063)
Pr_pipedwater_comm			0.087	0.221***			0.049	0.198**
			(0.089)	(0.083)			(0.089)	(0.084)
ISIC 3 controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Regional controls	Province	District	Province	District	Province	District	Province	District
Clustering	Village	Village	Commune	Commune	Village	Village	Commune	Commune
R-squared	0.376	0.394	0.372	0.390	0.570	0.581	0.569	0.580
n	484,567	484,567	491,773	491,773	493,191	493,191	500,582	500,582

Notes: Clustered standard errors presented in parenthesis. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.

Source: authors' calculations using the Cambodian Economic Census 2011.

References

Amiti, M., and Konings, J. (2007). 'Trade Liberalization, Intermediate Inputs and Productivity: Evidence from Indonesia'. *American Economic Review*, 97(5), 1611-38.

Audretsch, D., and Feldman, M. (1996). 'R&D Spillovers and the Geography of Innovation and Production'. *American Economic Review*, 86(3), 630-40.

Cortright, J. (2006) 'Making Sense of Clusters: Regional Competitiveness and Economics Development'. A Discussion Paper Prepared for the Brookings Institution Metropolitan Policy Program.

Crespo, N., and Fontoura, M. (2007). 'Determinant Factors of FDI Spillovers – What do we Really Know?' *World Development*, 35(3), 410-25.

Duranton, G., and Overman, H. (2005). 'Testing for Localization using Micro-Geographic Data'. *Review of Economic Studies*, 72(4), 1077-1106.

Ellison, G., and Glaeser, E. (1999) 'The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?' *American Economic Review*, 89(2), 311-16.

Fafchamps, M., and Söderbom, M. (2014). 'Network Proximity and Business Practices in African Manufacturing'. *World Bank Economic Review*, 28(1), 99-129.

Fan, C., and Scott, A. (2003). 'Industrial Agglomeration and Development: A Survey of Spatial Economic Issues in East Asia and a Statistical Analysis of Chinese Regions'. *Economic Geography*, 79(3), 295-319.

Fujita, M., Krugman, P., and Venables, J. (1999). *The Spatial Economy: Cities, Regions and International Trade*. Cambridge, MA: MIT Press.

Howard, E., Newman, C., and Thijssen, J. (2011). 'Are Spatial Networks of Firms Random? Evidence from Vietnam'. Working Paper 2011/87. Helsinki: UNU-WIDER.

Krugman, P. (1991). *Geography and Trade*. Cambridge, MA: MIT Press.

Manski, C. (1993). 'Identification of Endogenous Social Effects: The Reflection Problem'. *Review of Economic Studies*, 60, 531-42.

Marin, A., and Bell, M. (2006). 'Technology Spillovers from Foreign Direct Investment (FDI): The Active Role of MNC Subsidiaries in Argentina in the 1990s'. *Journal of Development Studies*, 42(4), 678-97.

Marshall, A. (1920). *Principles of Economics*. London: MacMillan.

McKenzie, D., and Sakho, Y. (2007). 'Does it Pay Firms to Register for Taxes? The Impact of Formality on Firm Profitability'. Policy Research Working Paper 4449. Washington, DC: World Bank.

Moreno-Monroy, A. (2012). 'Critical Commentary. Informality in Space: Understanding Agglomeration Economies during Economic Development'. *Urban Studies*, 49(10), 2019-30.

Overman, H., and Venables, T. (2005). 'Cities in the Developing World'. Discussion Paper 695. London: Centre for Economic Performance, London School of Economics and Political Science.

Rand, J., and Torm, N. (2012). 'The Benefits of Formalization: Evidence from Vietnamese Manufacturing SMEs'. *World Development*, 40(5), 983-98.

Rosenthal, S., and Strange, W. (2004). 'Evidence on the Nature and Source of Agglomeration Economies'. In, V. Henderson and J.F. Thisse (eds) *Handbook of Urban and Regional Economics*. North Holland: Elsevier.

Siba, E., Söderbom, M., Bigsten, A., and Gebreeyesus, M. (2012). 'Enterprise Agglomeration, Output Prices and Physical Productivity: Evidence from Ethiopia'. Working Paper 2012/85. Helsinki: UNU-WIDER.