Maryiam Haroon and Azam Chaudhry (2014). Where Do New Firms Locate? The Effects of Agglomeration on the Formation and Scale of Operations of New Firms in Punjab. Economics Discussion Papers, No 2014-21, Kiel Institute for the World Economy. http://www.economics-ejournal.org/economics/discussionpapers/2014-21

Responses to Comment:

Comment 1: Explain why and how you use the DOI 2010 to measure the arrival of firms in 2008 only. Why not for 2010? If that is possible, you can also make a dataset of arrivals for each individual year? In addition, does arrival mean a firm coming from somewhere else or is it a startup firm? Also interesting is if it is an independent firm. Maybe many of these arrivals could be branches or establishments of the same larger company (would you consider every opening of a new McDonald as an arrival of a new firm?).
The paper does not mention how many districts appear in the dataset which makes it unclear to interpret the results and degrees of freedom of the model. Please add a table with the nr of observations by district.

We use the data set of 2010 to determine new firm entry in 2008 since the dataset includes information about the year of establishment of each firm. So those firms whose year of establishment is 2008 have been classified as new arrival in 2008. We have not used 2010 as the year of analysis since we do not have the data for socio economic characteristics of the Punjabi districts for 2010 (and other time periods). The data set does not distinguish between new firms or movement of firms within an area which is one limitation of the dataset. It is possible that a firm is a branch of a large company, but we are primarily interested in where firms chose to locate. Even if a firm is a branch of large company we want to see where the large company decides to establish new branches. The data set includes 34 districts according to the classification and we have 1088 observations for the analysis which are industry- district combinations.

Comment 2: Equation (11) and (12) follow quite nicely from the literature discussed on page 1-4. Again this indicates that section 3 as a whole can better be deleted from the paper because it is not related in any way to the empirical specification.

The Soubeyran and Thisse (2008) model shows that a higher number of workers and higher knowledge spillovers are the two factors that attract new firms to locate in a specific area. So, the equations in our model are based to the number of workers. We subdivide the number of workers as in Soubeyran and Thisse (2008) into two components: the total number of workers in a district and the number of workers in a particular industry within a particular district. The existing model can be adapted by adding the superscript on the number of workers (L_d) in equation 10. The superscript will vary in two ways: it will vary across the number of workers in a region and number

of workers in an industry within a district. We moved from equation 10 to 11 and 12, since equation 10 states that higher number of workers will lead more new firms to enter in an area, which indicates that more employment will lead more new firms to enter. So, the number of workers (L_d) in equation 10 has been decomposed into two components in equation 11 and 12 and the two components are the number of in a particular industry in a district (which is referred to as localization) and the total employment in a district (which is referred to as urbanization).

Comment 3: What are "sub provincial effects" and why are they added? Is this something else than district? How many sub provinces are there?

A province can be divided into sub provincial regions and further into districts. There are four sub provincial regions in Punjab which are central Punjab, southern Punjab, northern Punjab and western Punjab (so a sub province is wider division than a district). The reason for including sub provincial effects is that there may be common characteristics at the sub- provincial level (so districts in southern Punjab share many of the same socio-economic characteristics) and in order account for any variation that exists at sub-provincial region level we have also controlled for sub-provincial effects.

Comment 4: For a measure of localization for industry *i* in district *d*, simply total employment have been taken. This is not localization but just size. Your model then predicts that more arrivals will happen in large districts. Instead, you need a localization indicator λ_d like for example the regional component of the sector share: local share of employment in – say – the footwear sector divided by the same at the national/provincial level). Then, if $\lambda_d > 1$, district *d* is relatively more specialized in footwear. See Combes, Mayer & Thisse (2008) chapter 10 for an elaborate discussion on localization and specialization measures. For a measure of urbanization the same problem appears. Here simply total employment is taken which again is only measuring size. Density, or the share of population of the largest city(ies) in the district could be taken.

There are several measures of localization and urbanization that have been used in the literature and we have chosen the employment level in order to measure localization and urbanization. Other papers have also constructed these measures by using employment levels (Rosenthal and Strange, 2010). We want to test how the presence of diverse activity is likely to attract new firms and density or share of population measures can be used to account for total demand in an area which we are not testing for in our model.

Comment 5: K_d is introduced as the initial level of knowledge in district *d* (line 6). One sentence further (line 7) the same symbol K_d is introduced as capital for which interest rate *r*

is charged. In the latter I assume it should read *k* assuming that this *k* is the same in every district?

The reviewer has pointed out a misstatement in the paper. In our paper, capital is represented as k and the sentence should be "Entrepreneurs can start a new firm by acquiring capital k at interest rate *r*, hiring labor in a particular district and can sell its product at price *p".* The interest rate is assumed to be the same in all the districts since the districts are small in comparison with rest of the world. The capital requirement of the firms is also assumed to the same across districts.

Comment 6: Further confusing come in when later it is said that capital *k(qd)* "is constant across districts" but apparently it is not fixed costs but only variable costs depend on output.

It is assumed that capital required by each firm depends upon the output it produces and each firm produces the same level of output hence the capital requirement by each firms is the same. Furthermore, it is assumed that the firms produce the same level of output across districts so capital requirement by the firms is also the same across districts.

Comment 7: In (1) symbol *l* is introduced as "worker's knowledge base", but then, what is *l(Kd)*? The worker's knowledge base as a function of the initial level of knowledge in district *d*? And why is it multiplied with output *qd* and the wage *wd*, apparently to represent labor cost? I suspect that *l* should be *ld* here meaning the number of workers hired in district *d* which is a negative function of *Kd* , presumably because with a higher knowledge base you need less labor? But then it should read *l(Kd ,qd)* because, just as *k(qd)*, the required number of workers must be some function of output.

Equation 1 represents the cost function of a firm in a district and the first part of equation (1) is the cost of labor. It is assumed that the cost of labor given in the equation (1) depends upon specificity of the district. It is because of this that wages and the capital of skills $l(K_d)$ are different across districts. The capital of skills in equation (1) can be regarded as workers with knowledge which have been accumulated through learning by doing and are independent of current output level of the firm. It only depends upon the knowledge base of workers.

Comment 8: From equation (2) it is clear that maximizing profits is the same as minimizing costs. These costs go down with K_d by definition because *l'<0* by definition. Therefore equation (3) is confusing and can be deleted. Apart from that, define what you mean with the hat (^) symbol here.

Equation (3) shows that the maximum profits a firm can derive which can be represented as ($\hat{\Pi}$) and this has been shown when we are maximizing the profit function.

Comment 9: Equation (4) is confusing: what is *K'* (the first derivative to what?). Equation (5) is self-evident and can be deleted. Equation (6) does not tell what λ is.

We maximize profits with respect to output and then by mathematical manipulation derived the profit maximizing level of output. The expression λ can be explained as follows:
$\lambda(q) = k'(q)q - k(q)$

Comment 10: Equation (7) should be deleted because defining the number of firms in (8) as total labor force divided by labor demand of one firm is clear enough. What is not clear, however, is why every firm is of the same size. This is an implicit assumption in (8) and needs to be addressed. "Output produced by firms is the same across districts" sounds like every district is of the same size, you probably mean identical firm size across all districts, which as far as I can see cannot be derived here but is an assumption.

Using equation 6 and the equality of profits between districts imply that $r\lambda(\hat{q}_d) = r\lambda(\hat{q}_e)$ where *d*, $e \in I$ and *I* represent districts where new firms will be established. This indicates that the firm's equilibrium output is the same. As all the firms in a district face the same environment thus the output across firms in a district is equal in equilibrium and we do not distinguish among the firms in a district. Also the output produced by firms across district is also assumed to be same. The district output level can vary as there is variation across number of firms in the district.

Comment 11: In equation (9) what is (hat) *q(I)* ? Your previous notation to indicate output by district is *qd*. Better use *Q* to indicate total output of all districts together which I believe you mean here. The symbol ν is said to be increasing here but it does not tell what it is. Better just delete (9) and say that you assume all labor supply is used in each region. This directly gives the number of firms in each district to be linear proportional to labor supply by district.

In equation (9), $q\hat{}(I)$ is the output produced by each firm in a district and *v* is the functional form of *K*. This has been used in order simplify the expression and we can state it as $v(K) = 1/l(K)$.

Comment 12: In the end, the only qualification that comes out of page 5-6 in equation (10) is that *nd* will be less that proportional to *Ld* because a higher stock of knowledge means more productive labor. For that, you might even consider to delete all equations because they do not add to any further understanding (redundant math exposure). Some more theory (and maybe equations) is needed here, however, to justify the assumption of all firms being

of the same size. This is normally an outcome of monopolistic competition models, which I expect Soubeyran & Thisse are also using.

The model has been taken from Soubeyran & Thisse and we have used in order to show how our empirical specification can be derived from theoretical model which has already been developed in the literature. The model developed in this study takes all the assumptions developed by Soubeyran & Thisse in order to derive the equations and these assumptions can be made clearer in the paper.