

《出口集聚、企业相关生产能力与企业出口扩展》

评阅书及作者修改说明（上）

发表信息

郭琪, 周沂, 贺灿飞. 出口集聚、企业相关生产能力与企业出口扩展[J]. 中国工业经济, 2020, (5): 137-155.

外审意见 1 及作者修改说明

谢谢审稿专家对文章选题和内容创新性的肯定, 您提出的文章主要问题以及具体修改建议显著提升了文章整体质量, 十分感谢! 下文是我们对您提出的建议给予的具体回复。

1. 选题意义评价

论文的选题很好。论文实证研究了区域出口集聚对企业出口产品扩展的影响, 以及区域出口集聚和企业相关产品生产能力交互作用对企业产品扩展的影响。论文发现 (1) 区域同产品出口集聚和相关产品出口集聚提高 1%, 企业出口该产品上升 0.31% 和 0.01%。也就是说, 同产品出口集聚对企业出口该产品的概率的影响是相关产品出口集聚影响的 30 倍。(2) 给定区域出口集聚水平下, 相关产品生产能力更高的企业出口该产品的概率越高。(3) 专业劳动力和中间产品共享 (供给侧)、市场信息共享 (需求侧) 都是区域出口集聚效应的影响机制。

2. 内容创新评价

论文的创新性还不错。论文的主要创新处在于区分同产品出口集聚和相关产品出口集聚对企业在该产品首次出口概率的影响。但是在实证方法仅采用线性概率模型是该论文的主要问题。因为出口是少数现象, 不管从产品还是企业层面而言, 所以 Logit 或者 Probit 模型是更为合适的选择。

3. 存在的主要问题 和具体修改建议

(1) 使用 Logit 或 Probit 模型。修改论文中公式 6 所示的计量模型。

回复: 根据中心极限定理(Central Limit Theorems), 大量相互独立的随机变量, 在采样次数足够大的时候 (一般要超过 30 次以上), 其均值或者和的分布也接近正态分布。本轮修改中, 模型样本量达到了上亿条。故而, 即使因变量为二元变量, 在大样本下, 其残差分布也趋于正态分布, 此时采用 LPM 来对模型进行估计也是无偏的 (Angrist and Pischke, 2008)。由于本文在估计过程中加入了企业、产品和时间的固定效应, 在 stata 中采用 Logit 和 Probit 模型来估计效率极低, 甚至用大型计算机也很难算出结果, 因此文章仍然采用 LPM 来进行估计。

(2) 对产品相关度指标作更多的统计性描述。论文交代了如何构建该指标, 但是没有对该指标的统计性描述。对这一指标的统计性描述, 作者可以借鉴在论文中引用的 Hidalgo et al. (2007)。

回复: 谢谢您的建议。听取您的建议, 文章已经借鉴 Hidalgo et al.(2007)对产品相关度指标做了简要的统计性描述以及网络图展示, 使读者和相关研究者对产品相关度有更清晰的认识 (具体的修改和调整情况见文章第三部分)。

图 1 展示产品相关度的核密度分布。产品相关度的分布严重左偏, 说明产品空间中的相关度大多比较弱, 这与很多研究结果高度一致 (Hidalgo et al., 2007; Boschma et al., 2012; Neffke et al., 2011)。大约有 1% 的关联在 0.32 以上, 产品相关度的平均值仅有

0.064, 超过 95%的关联在 0.2 以下, 这些弱关联不能说明两个产品之间有真正意义上的技术相关。本文借助 Cytoscape3.2.1 绘制中国 2002-2011 年期间的出口产品技术相关网络图, 见图 2。为了使该网络具有可读性, 绘制该图时只保留 0.3 以上的技术相关产品对, 包括 534 个 HS 四位码产品 (点) 和 8077 条关联 (边)。为了识别产品类别, 本文按照投入产出表产业大类将产品划分为 11 类, 并用不同颜色加以区分, 具体见图例。从图中可以发现, 大部分产品之间存在一个关联网络, 见图中③④⑤部分, 有一个明显的关联核心, 金属冶炼及制品业 (黑色)、部分化学工业 (玫瑰红)、电气机械及电子通信设备制造业 (深红色) 等产业位于这个网络的核心, 围绕在核心周围存在很多关联束, 这些关联束表现出明显的专业化特征, 即从核心向外延伸的每个方向主要以一种或两种产品为主, 例如, 左上部分③以机械工业为主, 右上部分一条以金属冶炼及制品业为主, 一条以化学工业为主, 右下部分⑤以纺织服装业为主, 左下部分④以电气机械及电子通信设备制造业为主, 这种同行业产品相互临近说明相同行业内的产品技术相关度更高。除核心结构外, 还存在很多零散破碎的关联, 如图 2 中①②⑥所示, 它们也表现出同行业产品相互临近的特征。

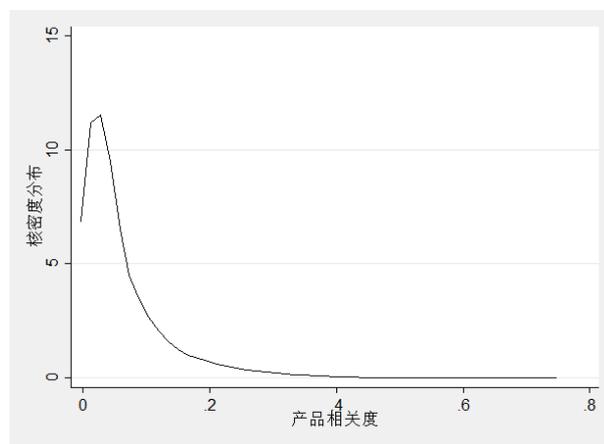


图 1 产品相关度 ϕ_{ij} 的核密度分布图

资料来源: 作者根据中国海关数据库 (2002-2011) 计算。

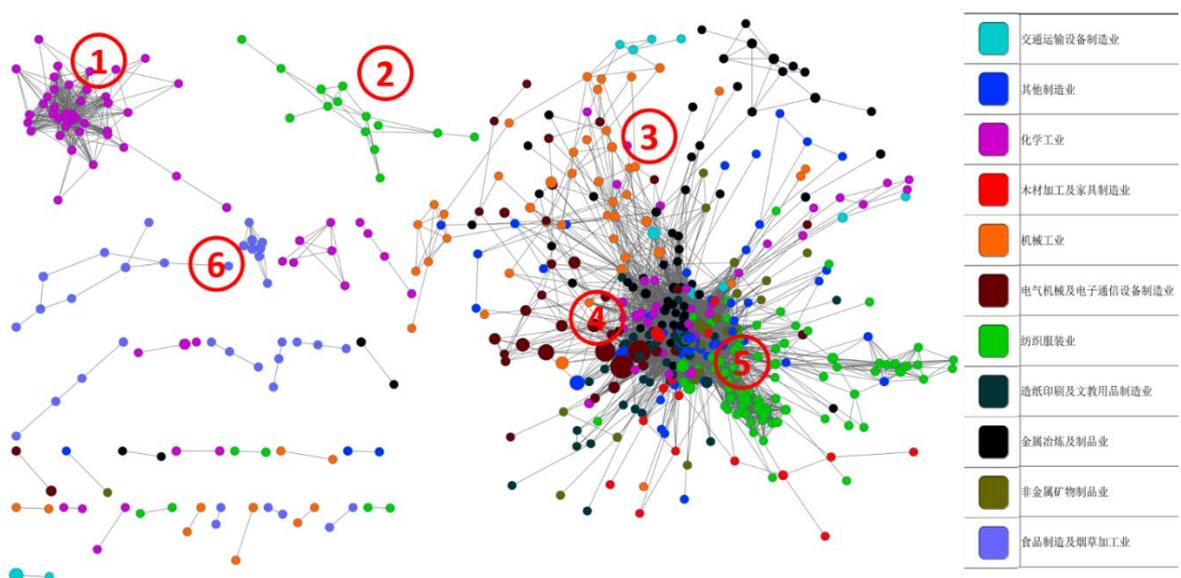


图 2 四位数 HS 码出口产品相关网络图

注：每个点代表不同的四位数 HS 码产品，不同颜色代表不同的大类产业，该分类依据区域间投入产出表的产业分类，具体对应见图例；点的大小代表 2002-2011 年不同产品出口额的大小，点越大出口额越大；该图用 cytoscape3.2.1 绘制而成，layout 选择 edge-weighted spring embedded layout。

资料来源：作者根据中国海关数据库（2002-2011）计算

（3）文章没有明确指出是使用一年的数据来计算产品相关度指标，还是用多年的数据计算的。不同年份的数据计算出的产品相关度指标有何差异吗？

回复：本文使用研究期（2002-2011 年）全样本企业数据计算产品相关度指标，有两个原因：一方面，使用全样本企业（504500 家非贸易公司），通过计算一对产品被同一企业出口的条件概率来反映产品的技术关联程度，以最大限度降低使用某一年数据所带来的偏误；另一方面，2002-2011 的十年是中国出口贸易发展最快的十年，产品种类变化较大，用任何一年都不足以反应这一时期中国产品相关度的全貌。考虑以上原因，本文采用 2002-2011 年多年数据计算。当然我们分别使用各年数据尝试计算了产品相关度，各年的产品相关度之间有非常高的相关系数，说明使用此方法估计的产品相关度是稳健的。

（4）区域溢出效应的指标的描述有不清楚的地方。在公式 2 中是对 j 加总，但是加总对象 $export_{c,i}$ 跟 j 没有关系。在公式 3,4 中，没有看明白为什么需要有减去 $\sum_j (export_{f,j} \times \delta_{ij})$ 这一项。

回复：感谢审稿人的意见。在公式 2 中应对 i 加总，作者笔误，实在抱歉。在公式 3,4 中，溢出效应的定义应为企业 f 所在城市 c 的其他企业在产品层面的集聚，因此，三类溢出效应均应剔除企业 f 本身在相应产品的出口规模，即 $\sum_j (export_{f,j} \times \delta_{ij})$ 这一项。同时，我们也对该指标的说明进行了修改和补充。

出口集聚效应的测度如下：

在获得“产品-产品”间的技术相关度后，本文计算区域溢出效应和企业生产能力。值得指出的是，本文所指的区域是城市。其中，区域溢出效应的测度中，根据产品之间认知距离的大小，又分为区域同产品溢出效应（ Agg ）、相关产品溢出效应（ Agg_R ）和不相关产品溢出效应（ Agg_UR ）。

区域同产品溢出效应（ Agg ）：

$$Agg_{c,i} = \sum_j (export_{c,i}) \quad (2)$$

其中， c 是企业 f 所在城市， i 和 j 为产品。 $export_{c,i}$ 为城市 c 出口产品 i 的规模。下同。溢出效应的定义应为企业 f 所在城市 c 的其他企业在产品层面的集聚，因此三类溢出效应均应扣除企业 f 本身在相应产品的出口规模。由于企业 f 在前一年未出口过产品 i ，因此企业本身对该产品的出口规模为 0。

区域相关产品溢出效应（ Agg_R ）：

$$Agg_R_{c,i} = \sum_j (export_{c,j} \times \delta_{ij}) - \sum_j (export_{f,j} \times \delta_{ij}), \text{ 其中, } \delta_{ij} \begin{cases} \delta_{ij} = 1, \text{ if } \phi_{ij} \geq 0.25 \\ \delta_{ij} = 0, \text{ if } \phi_{ij} < 0.25 \end{cases} \quad (3)$$

其中， ϕ_{ij} 为产品 i 和 j 的认知距离。为区分临近企业其他产品出口带来的溢出效应中相关

产品和不相关产品的作用，本文选择以 0.25¹作为产品间认知边界的临界值，产品技术相关度大于 0.25，则认为两产品在技术上相关，以此加权得到区域相关产品溢出效应。为了更为准确地衡量企业 f 所在城市 c 其他临近企业的溢出效应，区域相关产品溢出大小扣除了企业 f 本身相关产品出口的影响。

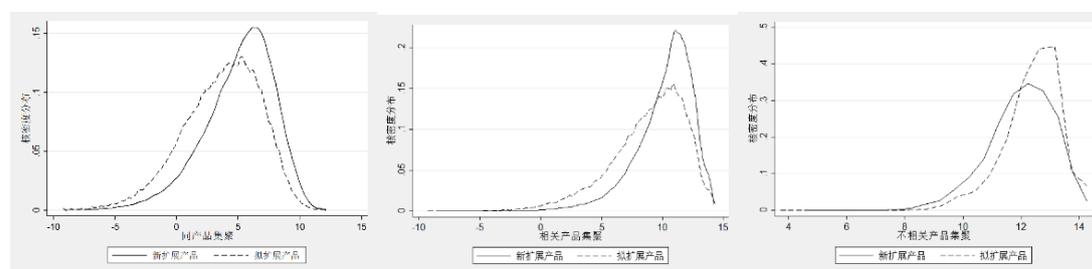
同理，区域不相关产品溢出 (Agg_UR) 如下：

$$Agg_UR_{c,i} = \sum_j (export_{c,j} \times \delta_{ij}) - \sum_j (export_{f,j} \times \delta_{ij}), \text{ 其中, } \delta_{ij} \begin{cases} \delta_{ij} = 0, \text{ if } \phi_{ij} \geq 0.25 \\ \delta_{ij} = 1, \text{ if } \phi_{ij} < 0.25 \end{cases} \quad (4)$$

(5) 对区域溢出效应的指标作更多的统计性描述。作者可以考虑借鉴 Heiwai Tang 和 Ana Fernandes 的文章 [Learning to Export from Neighbors](#)。

回复：感谢您的建议。由于本文区域溢出效应是在城市-产品层面的测算，对这一指标单独做统计性描述有难度，但为了初步刻画区域溢出效应与企业出口产品扩展之间的关系，我们做了更多的统计性描述，请见“三、数据描述与指标测算”的“4.特征性事实”部分。

为了初步验证三种类型出口集聚与企业出口产品扩展之间的关系，这里考察新扩展产品与拟扩展产品的出口集聚核密度分布差异，见图 3。其中，横轴分别是企业所在城市的同产品集聚、相关产品集聚和不相关产品集聚，均做取对数处理。结果与预期一致，对于同产品集聚（见图 3 (a)）和相关产品集聚（见图 3 (b)），拟扩展产品的分布在新产品的左侧，说明同产品集聚和相关产品集聚程度越高，该产品被企业扩展的概率更大。经 ANOVA 检验，p 值为 0.0000，即两个分布的差异在统计上显著。然而，不相关产品集聚（图 3 (c)）的作用恰恰相反，拟扩展产品的分布在新产品的右侧，说明不相关产品集聚程度越高，该产品被企业扩展的概率越小，初步证实不相关产品集聚效应可能表现为集聚不经济。为验证三种类型出口集聚对企业产品扩展的影响，下文将通过计量模型，在控制不同固定效应以及控制变量的基础上进一步讨论。



(a) 同产品集聚 (b) 相关产品集聚 (c) 不相关产品集聚

图 3 出口集聚与企业出口产品扩展的核密度分布

(6) 在样本数量上，论文有交代不清的地方。根据论文构建样本的方式，样本数量至少为，企业数量*年份数*随机抽取产品数=14687*10*50=7343500。但是论文中报告的样本量只有 5569787。论文应对此解释清楚。另外，表二中模型(3)和(4)用到的样本量少于表中其他模型，作者也应提供解释。

回复：本文的样本数量不会是严格的 14687*10*50=7343500，原因如下：第一，不是每个企

¹ 在总产品空间中，总产品数（点的数量）为 1084，总关联数（边的数量）为 587528，当门槛值设定为 0.1 时，所有产品仍都在产品空间中，但关联数仅剩五分之一，即 117685。当门槛值设定为 0.2 时，产品数剩下 80%，即 822，关联数为 32330；当门槛值设定为 0.3 时，产品数剩下 50%，即 534，关联数为 8077；当门槛值设定为 0.5 时，产品数仅剩约 10%，关联数为 173。门槛值的选择标准既要尽量包括更多的产品和关联，又要去掉那些不能代表有技术关联的弱联系，综上几种考虑，借鉴 Boschma et al.(2012)和 Neffke et al., (2011)均选择了 0.25，这里也将门槛值设定为 0.25，此时产品空间中有 660 个产品、16519 个关联对。后文也将使用 0.2 和 0.3 作为门槛值来做稳健性检验。

业在每一年都有扩展新产品,本文只考察企业有新扩展产品的年份;第二,研究期为10年,但由于解释变量均滞后1年,IV滞后2年,导致样本量也有相应的损失。因此,最终参与估计的样本数量应该是远远少于7343500这个数字。需要特别提及的是,在本轮修改中,本文放弃对拟扩展产品的抽样,直接采用全样本,所以拟扩展产品量以及总样本量有很大的提升。

参考文献:

- (1) Angrist, J. D. and J. S. Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 2008.
- (2) Hidalgo C. A., B. Klinger, and A. L. Barabási. The Product Space Conditions the Development of Nations[J]. *Science*, 2007, 317(5837): 482-487.

外审意见2及作者修改说明

感谢您对该文选题意义和创新性给予的充分肯定,您提出的文章主要问题以及具体修改建议显著提升了文章整体质量,十分感谢!下文是我们对您提出的建议给予的具体回复。

1. 选题意义

选题具有比较强的现实意义。特别是对于企业相关生产能力培育与企业出口拓展方面展开了一定的研究,将为产业政策的主力点选择或者偏重方向确定提供一定的参考。

2. 内容创新评价

(1) 本研究在出口溢出的基础上,将产品间的认知距离纳入到了引入到集聚经济分析中,进而研究了不同类型的区域溢出效应对企业产品扩展边际的影响,并对企业储备的相关生产能力的异质性与出口溢出研究做了一次有益的探索性研究。

(2) 在“出口集聚的影响机制检验”部分分析了企业的产品扩展方向。

3. 存在的主要问题

(1) 研究结论中,提到的“区域同产品出口集聚和相关产品出口集聚提高1%,企业出口扩展的概率分别上升0.31%和0.01%”,在此采用概率上升来说明,是否能够有力支持企业相关生产能力提升,从而增强企业出口拓展,是否还需要一些总量上的统计分析或者比例上的统计分析来更好地说明这一点研究。以及与“不相关产品出口集聚表现为竞争效应”这一结论做到有更为明确的说明。

回复:感谢您的建议。我们根据论文实际情况,在第三部分增加了“4.特征性事实”,以初步刻画区域溢出效应与企业出口产品扩展之间的关系。同时,对不相关产品出口集聚表现的竞争效应做了简要说明。由于竞争效应不是本文核心论点,所以仍然采用弱化的策略,以凸显同产品集聚和相关产品集聚带来的正向溢出效应。

为了初步验证三种类型出口集聚与企业出口产品扩展之间的关系,这里考察新扩展产品与拟扩展产品的出口集聚核密度分布差异,见图1。其中,横轴分别是企业所在城市的同产品集聚、相关产品集聚和不相关产品集聚,均做取对数处理。结果与预期一致,对于同产品集聚(见图3(a))和相关产品集聚(见图3(b)),拟扩展产品的分布在新产品的左侧,说明同产品集聚和相关产品集聚程度越高,该产品被企业扩展的概率更大。经ANOVA检验,p值为0.0000,即两个分布的差异在统计上显著。然而,不相关产品集聚(图3(c))的作用恰恰相反,拟扩展产品的分布在新产品的右侧,说明不相关产品集聚程度越高,该产品被企业扩展的概率越小,初步证实不相关产品集聚效应可能表现为集聚不经济。为验证三种类型出口集聚对企业产品扩展的影响,下文将通过计量模型,在控制不同固定效应以及控制变量的基础上进一步讨论。

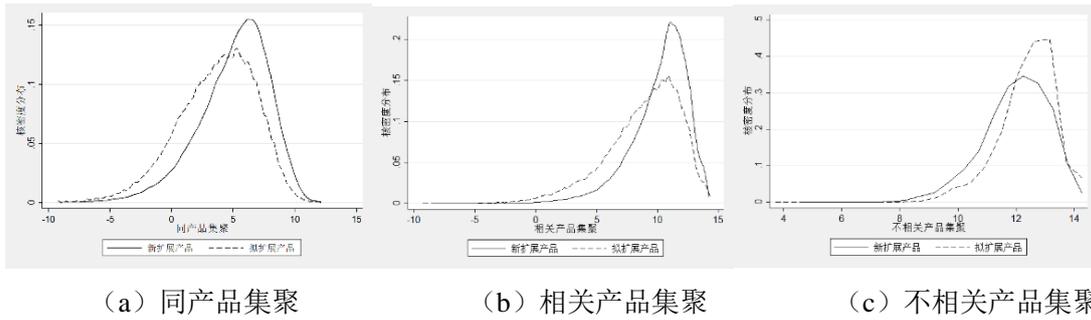


图 3 出口集聚与企业出口产品扩展的核密度分布

(2) 研究综述的逻辑顺序可能需要做一些基本的调整。在本部分中，作者在研究综述的开篇未能直接切入到本文的主要研究思维上。可能需要从一定程度上跳出现有的文献逻辑顺序，综合写出一段本研究是依托于哪些相关研究而展开，以及在逻辑上这些研究是如何为本研究的提出服务的。

回复：谢谢审稿人的意见。本次修改稿对研究综述和理论分析的逻辑顺序做了较大调整，在研究综述的开篇就直接切入主题，同时也精简了理论分析的部分，请参看第二部分。

(3) 第二部分中 P5 页“转换成本”一词，在本部分缺少一些必要的界定，这对如何提出假设二是一个基本的前提。现在的文稿中，对于两个假设的这部分叙述与文献综述的衔接做的不够明确。

回复：根据审稿人的意见，本文对“转换成本”部分做了调整，为了提出假说二，对于文献综述和假说提出之间加了部分理论分析的衔接，请参看第二部分相关内容。

(4) 对照公式 (2)，第 7 页公式 (3) 中的 $export_{c,j}$ 、 $export_{f,j}$ 的意义是否分别表示为城市 c 出口产品 j 的规模、企业 f 产品 j 的出口额，如果是，那么城市和企业纳入到了同一个公式中，是出于怎样的考虑，作者需要有一点基本的说明；而且对于公式 (3)、(4) 而言，其中的 f 只是一个企业，那么两个公式是不是表示为从城市 c 中 j 种产品出口的规模减去一家企业 j 种产品的出口额。可以说，在公式 (3)、(4) 中，对于下标 f 的考虑并不明确。

回复：谢谢审稿人的提醒。 $export_{c,j}$ 、 $export_{f,j}$ 分别代表城市 c 出口产品 j 的规模、企业 f 出口产品 j 的出口额。在公式 3,4 中，溢出效应的定义应为企业 f 所在城市 c 的其他企业在产品层面的集聚，因此三类溢出效应均应剔除企业 f 本身在相应产品的出口规模，即 $\sum_j (export_{f,j} \times \delta_{ij})$ 这一项，这也是为什么本文会将城市和企业纳入到同一个公式中来计算溢出效应。我们已对相应部分做出了修改和补充，请见文章第三部分。

在获得“产品-产品”间的技术相关度后，本文计算区域溢出效应和企业生产能力。值得指出的是，本文所指的区域是城市。其中，区域溢出效应的测度中，根据产品之间认知距离的大小，又分为区域同产品溢出效应 (Agg)、相关产品溢出效应 (Agg_R) 和不相关产品溢出效应 (Agg_UR)。

区域同产品溢出效应 (Agg):

$$Agg_{c,i} = \sum_j (export_{c,i}) \quad (2)$$

其中， c 是企业 f 所在城市， i 和 j 为产品。 $export_{c,i}$ 为城市 c 出口产品 i 的规模。下同。溢出效应的定义应为企业 f 所在城市 c 的其他企业在产品层面的集聚，因此三类溢出效应均应扣除企业 f 本身在相应产品的出口规模。由于企业 f 在前一年未出口过产品 i ，因此企业本身对该产品的出口规模为 0。

区域相关产品溢出效应 (Agg_R):

$$Agg_R_{c,i} = \sum_j(\text{export}_{c,j} \times \delta_{ij}) - \sum_j(\text{export}_{f,j} \times \delta_{ij}), \text{ 其中, } \delta_{ij} \begin{cases} \delta_{ij} = 1, \text{ if } \phi_{ij} \geq 0.25 \\ \delta_{ij} = 0, \text{ if } \phi_{ij} < 0.25 \end{cases} \quad (3)$$

其中, ϕ_{ij} 为产品 i 和 j 的认知距离。为区分临近企业其他产品出口带来的溢出效应中相关产品和不相关产品的作用, 本文选择以 0.25² 作为产品间认知边界的临界值, 产品技术相关度大于 0.25, 则认为两产品在技术上相关, 以此加权得到区域相关产品溢出效应。为了更为准确地衡量企业 f 所在城市 c 其他临近企业的溢出效应, 区域相关产品溢出大小扣除了企业 f 本身相关产品出口的影响。

同理, 区域不相关产品溢出 (Agg_UR) 如下:

$$Agg_UR_{c,i} = \sum_j(\text{export}_{c,j} \times \delta_{ij}) - \sum_j(\text{export}_{f,j} \times \delta_{ij}), \text{ 其中, } \delta_{ij} \begin{cases} \delta_{ij} = 0, \text{ if } \phi_{ij} \geq 0.25 \\ \delta_{ij} = 1, \text{ if } \phi_{ij} < 0.25 \end{cases} \quad (4)$$

(5) 第 8 页中提到“本文认为随机抽出的 50 个样本的分布可以代表拟扩展产品的总体分布, 抽样的拟扩展产品与企业扩展的新产品合并为本文核心模型的估计样本。”那么作者如何下的判断, 在这里是要给给出一些说明的。

回复: 这一点十分感谢审稿人。在我们重新审视抽样 50 的合理性时发现, 虽然抽样与全样本估计结果在统计显著性上没有差别, 但是二者在估计参数的经济意义上差别较大, 抽样不能完全代表全样本的估计结果。因此, 我们采用拟扩展产品的全样本作为估计样本, 并对文章中所有计量模型重新进行了估计, 结果更新请见文章。

(6) 表 2、表 4-表 8 中观测值的变化及相应说明不够清楚。特别是模型 (4) 和模型 (5)。此外, 在供给侧需求侧分析中, 一些结果的解释还需要更明确。

回复: 表 2 观测值的变化说明在文章中进一步明确: “这主要由于工业企业库和海关库的匹配过程中损失了大量观测值, 损失的观测值一方面是由于企业名称无法精确匹配, 另一重要方面是因为工业库中规模以下的中小企业样本缺失。而企业规模可能是影响企业出口扩展行为的重要因素, 故而中小企业样本的损失可能导致结果出现明显偏误。从表 2 第 (3) 列的系数来看, 集聚经济在规模以上企业样本中的系数更小, 这主要是因为由于内部规模经济的存在, 大企业在出口产品扩展的过程中可能对外部规模经济的依赖更小, 该结论也符合预期。为了避免匹配导致的观测值损失带来的偏误, 后文只保留来源于海关库的企业出口总额和企业出口产品多样化两个控制变量。为了证明来源于工业企业库的企业劳动生产率和企业规模两个控制变量对模型解释力和集聚的系数没有显著影响, 本文用匹配后的样本观测值仅控制出口库两个控制变量, 重新估计表 2 第 (3) 列, 结果见第 (4) 列, 对比 (3) 和 (4) 确实发现企业劳动生产率和企业规模这两个控制变量对结果几乎没有影响, 这一定程度上证明了本文只使用海关库更全样本观测值的合理性。”另外, 由于文章结构的调整, 表格编号已经变化。原表 4-8 中观测值变化是由于滞后变量作为 IV 进入模型、异质性和稳健性检验过程中将样本分类的原因。需要指出的是, 由于本文放弃对拟扩展产品进行抽样, 而将所有拟扩展产品均纳入样本, 所以修改稿中的样本数量与原稿相比有较大提高。

供给侧机制讨论如下:

如前文所述, 马歇尔外部性理论认为集聚效应来源于共享专业劳动力、中间投入品和技术溢出, 适用于从供给侧角度或者产品生产角度来解释出口集聚效应。出口集聚效应使“跟

² 在总产品空间中, 总产品数 (点的数量) 为 1084, 总关联数 (边的数量) 为 587528, 当阈值设定为 0.1 时, 所有产品仍都在产品空间中, 但关联数仅剩五分之一, 即 117685。当阈值设定为 0.2 时, 产品数剩下 80%, 即 822, 关联数为 32330; 当阈值设定为 0.3 时, 产品数剩下 50%, 即 534, 关联数为 8077; 当阈值设定为 0.5 时, 产品数仅剩约 10%, 关联数为 173。阈值的选择标准既要尽量包括更多的产品和关联, 又要去掉那些不能代表有技术关联的弱联系, 综上几种考虑, 借鉴 Boschma et al.(2012)和 Neffke et al., (2011)均选择了 0.25, 这里也将阈值设定为 0.25, 此时产品空间中有 660 个产品、16519 个关联对。后文也将使用 0.2 和 0.3 作为阈值来做稳健性检验。

随者”可以享受“先驱者”所积累下来的生产要素和技术，使其更高效地获取到更加匹配的劳动力、投入要素和生产技术，大大降低了企业出口产品扩展的沉没成本和创新风险。因此，本文将从专业劳动力共享、中间投入品共享以及技术溢出三个角度研究出口集聚效应的影响渠道，同时尝试探究企业相关生产能力通过哪种渠道影响企业在出口集聚中获益。

三个渠道的衡量方法如下：利用工业企业库，计算 c 城市 i 行业企业的年平均就业人数之和来测度专业劳动力共享 ($Labor_{ci}$)；参考韩峰和柯善咨 (2012) 的方法，利用 2002 年和 2007 年中国投入产出表计算 c 城市 i 行业对 j 行业的完全消耗系数³，加权各行业产出规模，得到 c 城市 i 行业的中间投入品共享 ($Input_{ci}$)；借鉴劳动力共享变量的计算方法，利用工业库中的企业新产品产值数据并在城市一行业层面加总，得到 c 城市 i 行业的技术溢出 ($Spillover_{ci}$)。三个变量均进行对数化处理。

这里采用中介效应方法检验出口集聚效应的供给侧渠道：①式 (1) 估计出口集聚对企业出口产品扩展的总效应；②式 (8) 检验出口集聚对三个中介变量的影响；③将出口集聚变量和中介变量同时放入模型，检验中介变量对出口集聚系数变化的影响，见式 (9)。三个方程的系数均显著才能说明存在中介效应，结果见表 2 和表 10。

$$M_{c,i,t} = \delta_0 + \delta_1 Agg_{c,i,t} + \delta_2 Agg_R_{c,i,t} + \delta_3 Agg_UR_{c,i,t} + \gamma_f + \gamma_t + \gamma_i + \varepsilon_{f,i,t} \quad (8)$$

$$Entry_{f,i,t} = \beta_0 + \alpha_0 M_{c,i,t-1} + \beta_1 Agg_{c,i,t-1} + \beta_2 Agg_R_{c,i,t-1} + \beta_3 Agg_UR_{c,i,t-1} + \gamma_f + \gamma_t + \gamma_i + \varepsilon_{f,i,t} \quad (9)$$

表 2 第 (1) 列显示，使用工具变量解决内生性问题后，出口集聚对企业出口产品扩展有显著的促进作用，同产品集聚程度提高 1%，企业向该产品扩展的概率提高 0.04%。表 10 报告出口集聚效应的渠道，第 (1) — (3) 列的结果显示出口集聚显著提高了当地的专业劳动力数量、中间投入品规模和技术研发水平，但第 (4) 列中只有专业劳动力、中间投入品规模对企业出口产品扩展有显著的促进作用，技术溢出反而降低了企业产品扩展的概率，且系数不显著。这一结果并不难理解，因为专业劳动力和中间投入品均是企业产品扩展所需要的投入要素，两者的本地化特征均较为明显。集聚带来的本地专业化劳动力市场和中间投入品规模大大降低了企业扩展产品过程中对劳动力、中间品的搜寻成本和运输成本。技术溢出没有表现出显著的正面中介作用，原因可能是技术研发投入较高表明该产品在当地拥有较高的技术实力，从侧面也可以反映该产品在当地存在较为激烈的竞争以及该产品可能具有较高的技术密集度。因此，从结果来看技术研发越高，企业向该产品扩展的难度越大。综合来看，中介效应的结果表明，出口集聚通过专业劳动力共享和中间投入品共享显著提高了企业出口产品扩展的概率。这一结论与孙天阳等 (2018) 一致，但不同的是，本文发现技术溢出不是出口集聚促进企业产品扩展的渠道。本文认为，中介效应的检验不应将不同作用渠道分别引入模型，因为不同作用渠道一般存在相互作用的可能。例如，研发水平较高的地区，其专业劳动力和中间投入品的规模也可能更大，因此，需要控制其他渠道的影响来识别出真正的目标渠道。

为了检验这三个渠道是否也影响企业相关生产能力在出口集聚效应中的作用，本文在式 (8) 基础上引入企业相关生产能力与三个作用渠道的交叉项进行检验，结果见表 10 第 (5) 列。专业劳动力与企业相关产品能力的交叉项显著为正，表明企业相关生产能力越高，越容易在出口集聚中享受到专业劳动力共享带来的正向外部性，从而有利于提高企业出口扩展的概率。技术溢出和中间投入品与企业相关能力的交叉项为负，表明企业相关能力越高，从当地获取技术溢出和中间投入品的概率反而越小。一方面，可能因为相关能力高的企业越容易实现内部规模经济，从而将技术研发和中间投入品的供给内部化；另一方面，这些企业也更有能力在更大空间范围内寻找合适的供应商，尤其是对运输距离不敏感的零配件。根据企业

³采用年份最近距离匹配，2002-2005 年完全消耗系数使用 2002 年投入产出表的计算结果，2006-2011 年使用 2007 年投入产出表的计算结果。

资源视角，企业储备的能力与某产品在技术上相关度越高，企业向该产品扩展需要投入的转换成本越低。企业相关生产能力提高了企业对技术与知识的转换能力，而不具备这些能力的企业虽然也可以享受出口集聚效应，但效率和成功率显著低于能力强的企业。更值得注意的是，企业相关生产能力越高，其越有能力将技术研发与中间投入内部化，此时反而会降低对集聚效应的依赖。

表 10 出口集聚的影响机制检验：供给侧

	(1) <i>Labor</i>	(2) <i>Input</i>	(3) <i>Spillover</i>	(4) <i>Entry</i>	(5) <i>Entry</i>
<i>Agg</i>	0.0532*** (5.29e-05)	0.0107*** (2.06e-05)	0.0839*** (0.0004)	0.0003*** (5.82e-06)	0.0003*** (5.81e-06)
<i>Labor</i>				0.0006*** (2.20e-05)	0.0005*** (2.20e-05)
<i>Input</i>				0.0022*** (5.44e-05)	0.0023*** (5.44e-05)
<i>Spillover</i>				-0.0000 (2.77e-06)	-0.0000* (2.77e-06)
<i>Labor</i> × <i>CAPAB</i>					0.0009*** (1.40e-05)
<i>Input</i> × <i>CAPAB</i>					-0.0018*** (1.83e-05)
<i>Spillover</i> × <i>CAPAB</i>					-8.92e-06*** (1.05e-06)
观测值	32688701	32859828	32688793	32688701	32688701
R^2	0.7240	0.9550	0.8270	0.0220	0.0220

注：所有模型均控制了 *Agg_U*、*Agg_UR*、*CAPAB*、*EXPORT* 以及企业、产品和年份固定效应。

需求侧机制讨论如下：

研发新产品和开拓新市场是企业出口扩展的两种选择，向新产品扩展需要投入生产技术、资本和劳动力的转换成本，向新市场扩展需要投入搜集目的国的需求量、消费者偏好和出口渠道等市场信息的成本。企业在出口扩展时需要同时考虑在产品 and 市场两个维度的成功率。因此，与一般集聚外部性不同的是，出口集聚的影响渠道除专业化劳动力共享、中间投入品规模以及技术外溢等供给侧维度以外，还需要考虑需求侧信息积累对企业出口扩展的影响。一般来说，如果 A 国是 *c* 城市 *i* 产品的主要出口目的国，那么 *c* 城市围绕 A 国所积累的市场信息将更有利于企业向 *i* 产品扩展。

为了考察目的国需求侧信息优势是否有利于促进企业出口扩展，机制检验模型因变量延伸为企业—产品—目的国维度，自变量延伸为城市—产品—目的国维度的集聚变量，即城市—同产品—目的国集聚 (*Aggc*)、城市—相关产品—目的国集聚 (*Aggc_R*)、城市—不相关产品—目的国集聚 (*Aggc_UR*)。模型中控制企业出口目的国经验、企业相关生产能力、企业出口规模和出口产品多样化水平等企业特征以及企业、产品、年份和目的国固定效应，结果见表 11。第 (1) 列显示，同产品—目的国集聚程度提高 1%，企业向该产品—目的国扩展的概率提高 0.07%；相关产品—目的国集聚程度提高 1%，企业向该产品—目的国扩展的概率提高 0.01%；而不相关产品—目的国集聚依然降低了企业产品扩展的概率。此外，企业过去对目的国的出口经验也会影响企业出口扩展行为。表 11 第 (2) 列加入三个集聚变量与企业在该目的国的出口经验（即如果企业出口过该目的国，则 *EXP* 为 1，否则为 0）的交叉项。结果

显示，企业如果对某国有过出口经验，而本地其他企业又同时大量向该国出口产品 i 或者与 i 相关度高的产品，那么该企业向该国出口产品 i 的概率更高。这主要是因为，企业掌握的需求侧信息大大降低了向该目的国进行产品扩展的风险和不确定性。由此可见，目的国维度的信息和渠道优势显著提高了企业扩展边际，而企业出口经验将影响其对需求侧信息的吸收，两者共同影响企业扩展边际。

表 11 出口集聚的影响机制检验：需求侧

	(1) <i>Entry</i>	(2) <i>Entry</i>
<i>Aggc</i>	0.0007*** (6.81e-06)	0.0006*** (1.64e-05)
<i>Aggc_R</i>	0.0001*** (1.98e-06)	-0.0002*** (5.15e-06)
<i>Aggc_UR</i>	-0.0004*** (4.90e-06)	-0.0002*** (1.23e-05)
<i>Aggc×EXP</i>		0.0002*** (1.78e-05)
<i>Aggc_R×EXP</i>		0.0003*** (5.36e-06)
<i>Aggc_UR×EXP</i>		-0.0002*** (1.32e-05)
观测值	105401982	105401982
R^2	0.0350	0.0350

注：所有模型均控制企业出口目的国经验、企业相关产品能力以及企业出口总额和企业多样化水平等企业特征以及企业、产品、年份和目的国的固定效应。

(7) 引言部分的结论性内容和创新启示性内容，可以在此处删去。一些专有名词第一次出现英文缩写时，未能标注清楚。如 p4 页 MAR。

回复：已经按照您的建议对引言部分做了修改，同时，全文统一改成“马歇尔外部性”。

4. 具体修改建议

(1) 如 P4 页的表述应作为一部分首先在文献综述中直接显现。进而重新梳理研究综述。

回复：已对文章综述和理论分析重新进行梳理，请参见文章第二部分。

(2) 文献综述适当压缩，调整逻辑结构。同时对于两个基本假设的提出前，要有一个明确的思维性的表达或者概括，再进而结合本文的研究思路提出这两个假设。

回复：已对文献综述适当压缩，并围绕两个基本假说调整了逻辑结构，并加入一些理论分析进而提出理论假说。

(3) 全文实证部分的各小部分的模型分析中，建议有一个整体的思考。目前的各小部分的分析依托了不同的模型，它们之间的整体关系如何，是怎么来刻画本研究的基本思路的。也需要统一模型编号，避免在文章同样的表达，如“模型（1）、模型（2）”，在文章不同部分，实际上却代表了不同的模型。

回复：感谢您的建议。对实证部分按照分析逻辑重新进行了结构上的调整，对于模型编号和相应分析也做了统一的修改。请参看文章实证分析的部分。

(4) 文章的整体逻辑性有待加强，在内容组织上，要紧密围绕着文章的基本思维而展开。

回复：感谢您的建议。本文对文章结构和整体逻辑性做出了较大修改，与原稿相比，在内容组织上更加围绕基本思路展开。

《出口集聚、企业相关生产能力与企业出口扩展》评阅书及作者 修改说明（下）

感谢编辑以及两位审稿人宝贵的修改意见，我们按照意见书逐条修改。十分感谢编辑部和审稿人的辛苦付出！具体对于外审意见书的回复请见下文。

外审意见 1 及作者修改说明

谢谢作者根据上一轮建议所做的修改。但是回复和修改仍有改进和不清楚的地方，我把具体意见列举如下：

1. 针对产品相关度的描述，请列举产品对销量之和前十并且产品相关度大于 0.3 的产品对，以及每个产品在产品对中的销量占比，以及具体的产品相关度值。类似于这样的描述能给读者一个直观的了解，销量高并且相关度高的产品对是哪些产品。

回复：谢谢，根据您的建议，文中增加表 A1 来展示 2002 年和 2011 年中国出口额之和排名前 10 且相关度大于 0.3 的产品对，可以发现中国十年间出口产品结构从纺织服装等轻工业为主转变为机械与设备制造业为主。2002 年排名前 10 的产品对均以针织或钩编的套头衫、开襟衫、外穿背心及类似品（6110）为中心，相应的关联产品主要是能够与其共享上游纺织产品、低成本劳动力、相似生产技术和相似出口市场的服装、鞋帽、箱包等其他轻工业产品，在图 A2 中体现为⑤，即纺织服装业关联束。2011 年则有较大变化，排名前 10 的产品对以电话机（8517）为中心，相应的关联产品主要是其上游的零部件产品，例如变压器、传声器、集成电路、电线以及其他零件，在图 A2 中体现为关联核心④，即电子通信设备制造业关联束。综上所述，本文基于共存分析方法计算的产品相关度指标与实际情况基本相吻合，且其基于产业细类的测算远远优于以往粗糙的测度方法。

表 A1 出口额之和排名前 10 且相关度大于 0.3 的产品对

2002 年				2011 年			
产品 1	产品 2	产品相关 度	产品对出口额之和 (单位: 百万美 元)	产品 1	产品 2	产品相关 度	产品对出口额之和 (单位: 百万美 元)
6110	6204	0.54	2788.49	8517	8525	0.43	33573.22
4203	6110	0.31	2721.73	8504	8517	0.41	29100.11
4202	6110	0.34	2438.40	8517	8544	0.34	28914.26
6110	6203	0.48	2347.37	8517	8542	0.33	28283.79
6109	6110	0.69	2339.90	8517	8518	0.44	27677.68
6104	6110	0.75	2328.82	8517	8529	0.50	26400.37
6110	6302	0.31	2309.04	8473	8517	0.46	25701.39
6110	6205	0.39	2055.77	8517	8543	0.41	25633.34
4203	6402	0.32	2053.97	8517	8538	0.31	25488.67
6103	6110	0.66	2047.23	8517	9013	0.32	25140.59

注：4202 是指衣箱、提箱、小手袋、公文箱、公文包、书包、眼镜盒、望远镜盒、照相机盒、乐器盒、枪套等；4203 是指皮革或再生皮革制的衣服及衣着附件；6103 是指针或钩织的男式西服套装、便服套装、上衣、长裤、护胸背带工装裤、马裤及短裤（游泳裤除外）；6104 是指针或钩织的女式西服套装、便服套装、上衣、连衣裙、裙子、裙裤、长裤、护胸背带工装裤、马裤及短裤（游泳裤除外）；6109 是指针织或钩编的 T 恤衫、汗衫及其他背心；6110 是指针织或钩编的套头衫、开襟衫、外穿背心及类似品；6203 是指男式西服套装、便服套装、上衣、长裤、护胸背带工装裤、马裤及短裤（游泳裤除外）；6204 是指女式西服套装、便服套装、上衣、连衣裙、裙子、裙裤、长裤、护胸背带工装裤、马裤及短裤（游泳裤除外）；6205 是指男衬衫；6302 是指床上、餐桌、盥洗及厨房用的织物制品；6402 是指橡胶或塑料外底及鞋面的其他鞋靴；8473 是指专用于或主要用于品目 8469

至 8472 所列机器的零件、附件（罩套、提箱及类似品除外）；8504 是指变压器、静止式变流器及电感器；8517 是指电话机；8518 是指传声器及其座架、扬声器、耳机、耳塞；8525 是指无线电广播、电视发送设备；8529 是指专用于或主要用于品目 8525 至 8528 所列装置或设备的零件；8538 是指专用于或主要用于品目 8535 至 8537 所列装置的零件；8542 是指集成电路；8543 是指该章其他品目未列明的具有独立功能的电气设备及装置；8544 是指绝缘电线、电缆及其他绝缘电导体；9013 是指其他品目未列明的液晶装置、激光器及光学仪器、器具。

2. 针对区域溢出效应等指标，修改后的论文仍有许多解释不清楚的地方。下面依次指出修改版论文中公式（2）（3）（4）以及对公式相应的解释部分不一致的地方。

公式（2）：①修改版论文对公式（2）的解释（公式（2）下面灰色突出的一段）意思是**区域同产品溢出效应(Agg)**是一个企业-城市-产品（fci）层面的指标—“溢出效应的定义应为企业 f 所在城市 c 的其他企业在产品层面的集聚，因此三类溢出效应均应扣除企业 f 本身在相应产品的出口规模”，但是根据公式（2）的定义该指标是一个城市-产品层面的指标。②修改后的公式 2 仍是对 j 加总，但是加总对象 $export_{c,i}$ 跟 j 没有关系。这种公式本身和其解释的前后不一致让读者非常困惑。③根据修改版论文对公式（2）的解释，我理解的公式（2）应该如下面的形式：

$$Agg_{f,c,i} = export_{c,i} - export_{f,c,i},$$

其中 $export_{c,i}$ 是城市 c 在产品 i 的总出口， $export_{f,c,i}$ 是在城市 c 的企业 f 在产品 i 上总出口。“三类溢出效应均应扣除企业 f 本身在相应产品的出口规模”。请作者予以回复确认，以及做相应的修改。

公式（3）：**区域相关产品溢出效应 (Agg_R)** 是一个企业-城市-产品（fci）层面的指标还是一个城市-产品（ci）层面的指标，请作者予以回复确认和修改，论文的解释和公式（3）有明显的**不一致**。

公式（4）：**区域不相关产品溢出 (Agg_UR)** 是一个企业-城市-产品（fci）层面的指标还是一个城市-产品（ci）层面的指标，请作者予以回复确认。在公式中不用重复使用 δ_{ij} 给一个新的定义，把 δ_{ij} 换成 $(1 - \delta_{ij})$ 即可。

回复：感谢您细致且中肯的建议，文中关于公式的标识的确出现错误，已经认真修改。具体：

（1）三个溢出效应的指标均是企业-城市-产品层面的指标，公式中已经修改；（2）公式 2 应对企业 f 加总，请见文章；（3）公式 3 和公式 4 在“扣除企业本身在相应产品的出口规模”也做了更清晰的说明；（4）把 δ_{ij} 换成 $(1 - \delta_{ij})$ 的确更为严谨。具体修改如下，公式编号略有变化。

在获得“产品—产品”间的技术相关度后，本文通过识别企业所在的城市，计算出口集聚和企业相关生产能力。其中，出口集聚的测度中，根据产品之间认知距离的大小，又分为同产品集聚（Agg）、相关产品集聚（Agg_R）和不相关产品集聚（Agg_UR）。同产品集聚程度（Agg）计算公式如下：

$$Agg_{c,g,i} = \sum_{f \in c, f \neq g} (export_{f,i}) \quad (4)$$

其中，c 是企业 f 和 g 所在的城市，i 为产品。 $export_{f,i}$ 为城市 c 中企业 f 出口产品 i 的规模。同产品集聚效应是指企业 g 所在城市 c 的其他企业出口产品 i 的集聚程度 $Agg_{c,g,i}$ 对企业出口扩展的影响。通过引入产品技术关联，将大于某一技术关联值的产品定义为具有技术关联的产品对，相关产品集聚程度（Agg_R）计算公式如下：

$$Agg_R_{c,g,i} = \sum_j (export_{c,j} \times \delta_{ij}) - \sum_j (export_{g,j} \times \delta_{ij}), \text{ 其中, } \delta_{ij} = \begin{cases} 1, & \emptyset_{ij} \geq 0.25 \\ 0, & \emptyset_{ij} < 0.25 \end{cases} \quad (5)$$

其中， $export_{c,j}$ 和 $export_{g,j}$ 分别为企业 g 所在城市 c 和企业 g 出口产品 j 的总规模。 \emptyset_{ij} 为基于公式（1）计算的产品 i 和 j 之间的技术相关度。为区分相关产品和不相关产品集聚，

本文选择以 0.25 作为产品间技术相关度的临界值⁴。 ϕ_{ij} 大于等于 0.25，则认为两产品在技术上相关， δ_{ij} 取 1，否则 δ_{ij} 取 0。以 δ_{ij} 加权分别得到城市 c 中与产品 i 技术相关产品的出口总规模 ($\sum_j(\text{export}_{cj} \times \delta_{ij})$) 和城市 c 中企业 g 与产品 i 技术相关产品的出口规模 ($\sum_j(\text{export}_{gj} \times \delta_{ij})$)，两者相减得到企业 g 所在城市其他企业与产品 i 技术相关产品的出口集聚程度 $\text{Agg}_R_{c,g,i}$ ，其对企业 g 出口产品扩展的影响即为相关产品集聚效应。同理，识别技术关联度小于 0.25 的产品对，以此计算不相关产品集聚程度 (Agg_{UR})，计算公式如下：

$$\text{Agg}_{UR}_{c,g,i} = \sum_j[\text{export}_{cj} \times (1 - \delta_{ij})] - \sum_j[\text{export}_{gj} \times (1 - \delta_{ij})],$$

$$\text{其中, } \delta_{ij} = \begin{cases} 1, \phi_{ij} \geq 0.25 \\ 0, \phi_{ij} < 0.25 \end{cases} \quad (6)$$

所有符号的含义同上。不相关产品集聚 ($\text{Agg}_{UR}_{c,g,i}$) 对企业 g 出口扩展的影响即为不相关产品集聚效应。

3. 针对区域溢出效应的指标作更多的统计性描述。作者可以考虑借鉴 Heiwai Tang 和 Ana Fernandes 的文章 [Learning to Export from Neighbors](#)。尽管论文的区域溢出指标是在城市-产品层面的，但是作者可以选取几个有代表性的产品，借鉴 Heiwai Tang 和 Ana Fernandes 的文章用地图的方式把区域溢出效应做一定的描述。

回复：此次修改借鉴 Heiwai Tang 和 Ana Fernandes 的文章，选取两个有代表性的产品，用地图将三类区域溢出效应展示出来，由于文章篇幅限制，放在附录中。

4. 作者对于样本数量的回复，仍有不清楚的地方。“第一，不是每个企业在每一年都有扩展新产品，本文只考察企业有新扩展产品的年份”——跟企业产品扩展 $\text{Entry}_{f,i,t}$ 的定义，企业没有在产品 i 上扩展，根据定义就是零。那么也自然应该把这情况作为一观测值放在回归样本当中。为什么只考察企业有新扩展产品的年份？“第二研究期为 10 年，但由于解释变量均滞后 1 年，IV 滞后 2 年，导致样本量也有相应的损失。”——以基准回归为例，解释变量之后一起的话，应该还有 9 年的样本，那么样本量应该是 $14687 \times 9 \times 50 = 6609150$ 。“在本轮修改中，本文放弃对拟扩展产品的抽样，直接采用全样本，所以拟扩展产品量以及总样本量有很大的提升。”——如果不对产品抽样的话，那么样本量应该是在一亿 ($14687 \times 9 \times 1000$) 的量级，而不是目前样本量三千万的量级。如果使用全样本回归应当加上企业没有在产品 i 上扩展的观测值。

回复：谢谢您的质疑。作者对样本量做了简单的统计，在十年间一直在出口的企业 14687，其中：（1）每年有产品扩展的企业数量平均约为 6000 家（见下面表 S1，未放入正文），需要说明的是，本文模型中之所以只包括企业有新扩展产品的年份是因为模型一旦控制了企业、年份固定效应，那些 Entry 取值全部为 0 的企业在模型估计过程中就会被删掉，不参与计算；（2）每年每个企业扩展产品和拟扩展产品的总数量平均约为 600 种左右（见下面表 S2，未放入正文），因为本文对于扩展产品和拟扩展产品的定义近乎苛刻，即“企业 g 在 $t-1$

⁴在产品相关度矩阵中，有 1084 种产品，587528 对关联数。当阈值设定为 0.1 时，所有产品都在产品相关网络中，但关联数剩五分之一，即 117685；当阈值设定为 0.2 时，产品剩 822 个，关联数剩 32330；当阈值设定为 0.3 时，产品剩 534，关联数剩 8077；当阈值设定为 0.5 时，产品数仅剩 173。阈值的选择标准既要尽量包括更多的产品和关联，又要去掉那些不能代表有技术关联的弱联系。综上，借鉴 Neffke et al. (2012) 以 0.25 为阈值的设置，本文也将阈值设为 0.25，此时产品空间中有 660 个产品、16519 对关联对。后文也将使用 0.2 和 0.3 作为阈值来进行稳健性检验。

年及其之前所有年份没有出口产品 i ，而在 t 年开始出口，则 $Entry_{g,i,t}$ 取值为 1，产品 i 则为企业 g 扩展的新产品；企业 g 在 t 年之前（包括 t 年）的所有年份未出口 i 产品，则 $Entry_{g,i,t}$ 取值为 0”。也就是说，企业 g 在 $t-1$ 年及其之前只要有一年出口过产品 i ，该样本就会被删掉；（3）根据以上统计，按照 9 年计算，样本量是 $9*6000*600=32,400,000$ ，几乎接近本文模型估计的样本量。

表 S1 每年有产品扩展的企业数量

年份	企业数量
2003	7144
2004	6776
2005	6572
2006	6304
2007	5913
2008	5335
2009	5030
2010	4724
2011	4170

表 S2

年份	企业层面平均产品数量
2003	577.6788
2004	596.9439
2005	623.8545
2006	648.9735
2007	664.549
2008	659.6988
2009	653.273
2010	651.2398
2011	654.6535

5. 小提示：论文基准回归公式（6）中， $Entry_{f,i,t}$ 应当是企业-城市-产品-年份层面指标，公式中缺少了城市下标。相应的残差下标也要做调整。

回复：谢谢您的提示，已经按照您的建议作出修改。

为验证假说 1，本文将出口集聚分解为相同产品集聚和不同产品集聚，后者基于技术相关度进一步分解为技术相关产品集聚和不相关产品集聚。为了考察三种类型的出口集聚对中国企业出口扩展边际的影响，本文建立计量模型如（1）式所示。

$$Entry_{c,g,i,t} = \alpha_0 + \beta_1 Agg_{c,g,i,t-1} + \beta_2 Agg R_{c,g,i,t-1} + \beta_3 Agg UR_{c,g,i,t-1} + Control_{c,g,i,t-1} + \gamma_{g,i} + \gamma_{c,i} + \varepsilon_{c,g,i,t} \quad (1)$$

其中，下标 c 、 g 、 i 、 t 分别表示区域、企业、产品和年份。 $Entry$ 是二元变量，表示企业是否扩展产品 i ，扩展为 1，否则为 0。 β_1 、 β_2 、 β_3 分别考察同产品集聚效应、相关产品集聚效应和不相关产品集聚效应，所有解释变量均做取对数处理。 $Control$ 为企业—产品层面或企业层面的控制变量，包括企业相关生产能力（ $CAPAB$ ）、企业出口总额（ $EXPORT$ ）、企业出口产品多样化水平（ DIV ）、企业劳动生产率（ $PROV$ ）和企业规模（ $SIZE$ ）。其中，企业出口产品多样化水平为企业出口产品的种类数，企业劳动生产率为单位劳动力的企业生产总值，企业规模

为企业从业人数，企业出口总额和企业规模做取对数处理。

在（1）式的基础上，本文引入企业相关生产能力 $CAPAB$ 与三个出口集聚变量的交互项，构建计量模型来验证假说 2，即企业相关生产能力是否能提高企业在出口扩展过程中从出口集聚中的获益程度，如（2）式所示。

$$Entry_{c,g,i,t} = \alpha_0 + \beta_1 Agg_{c,g,i,t-1} + \beta_2 Agg_{-c,g,i,t-1} + \beta_3 Agg_{-UR_{c,g,i,t-1}} + \beta_4 Agg_{c,g,i,t-1} \times CAPAB_{g,i,t-1} + \beta_5 Agg_{-c,g,i,t-1} \times CAPAB_{g,i,t-1} + \beta_6 Agg_{-UR_{c,g,i,t-1}} \times CAPAB_{g,i,t-1} + Control + \gamma_{g,i} + \gamma_{c,i} + \varepsilon_{c,g,i,t} \quad (2)$$

其中，所有符号含义与（1）式一致。交叉项系数 β_4 、 β_5 、 β_6 分别考察企业能力对同产品集聚效应、区域相关产品集聚效应和区域不相关产品集聚效应的影响。为避免交叉项间的多重共线性问题，三个交叉项将被分别引入（2）式。

企业出口产品扩展的定义通常分为两年定义法或多年定义法。由于企业出口产品变动频繁，为了规避企业过去出口的影响以及识别企业对该产品的首次出口，本文采用多年定义法。企业 g 在 $t-1$ 年及其之前所有年份没有出口产品 i ，而在 t 年开始出口，则 $Entry_{g,i,t}$ 取值为 1，产品 i 则为企业 g 的新扩展产品；企业 g 在 t 年之前（包括 t 年）的所有年份未出口 i 产品，则 $Entry_{g,i,t}$ 取值为 0，产品 i 则为企业 g 的拟扩展产品。

外审意见 2 及作者修改说明

1. 在“理论分析与假设的提出”的这一部分，其中的理论分析还需要提炼出来几点明确地表达，不是像现在这样偏向于文献综述或者带一点点述评的段落。我们需要的是作者在文献或者理论梳理的基础上，更为明确地提出本文研究的基本理论基础是什么或者作者是坚持了哪些理论点，并结合本文的研究思路，从而提出了对应的几个假设。

回复：您的建议对于文章提升质量有至关重要的作用，十分感谢！作者已经重新梳理了理论假设部分，将原本的文献综述写法改成了理论机制分析，请见修改稿第二部分“理论分析与研究假说的提出”。

2. 对图 2 的解释，个人建议在表达上做一点转换。先谈结论性的那半句话，然后指示出在图 2 中的哪个部分。现在的表达，先说位置感觉不是很妥当。给阅读者的感觉，这样的解释是带有了很强的位置感觉的。毕竟各产品关联度更多的是一般意义上的表达，并不带有位置信息的。

回复：根据您的建议，作者对图 2 的解释重新修改，先谈结论，最后标出位置。修改后此部分的可读性明显提高。

基于产品相关度，本文绘制 2002-2011 年期间中国出口产品技术相关网络图，见图 2。从图中可以发现，大部分产品位于一个紧密的关联网络中，有一个明显的关联核心，金属冶炼及制品业（黑色）、部分化学工业（玫瑰红）、电气机械及电子通信设备制造业（深红色）等产业位于关联网络的核心；围绕在核心周围存在很多关联束，这些关联束表现出明显的专业化特征，即从核心向外延伸的每个方向主要以一种或两种产品为主，这种同行业产品相互临近说明相同行业内的产品技术相关度更高，这些关联束包括机械工业关联束（见图 2 中③）、纺织服装业关联束（见图 2 中⑤）、电气机械及电子通信设备制造业关联束（见图 2 中④）；除核心结构外，还存在很多零散破碎的关联，也表现出同行业产品相互临近的特征，如图 2 中①②⑥所示。

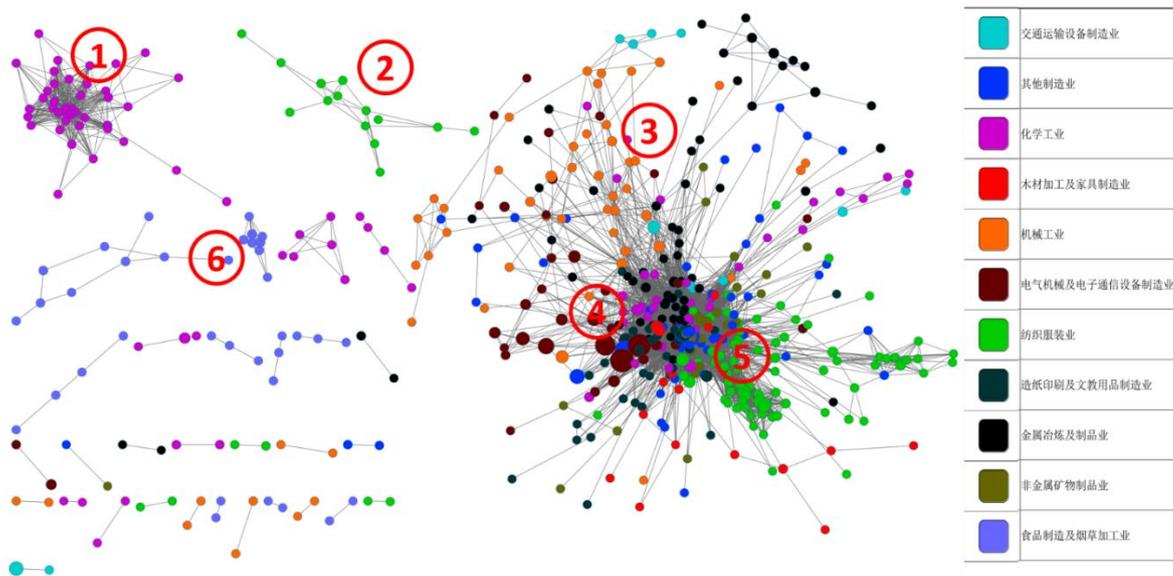


图 2 HS 四位码出口产品相关网络图

注：每个点代表不同的 HS 四位码产品，不同颜色代表不同的大类产业，该分类依据区域间投入产出表的产业分类，具体对应见图例；点的大小代表 2002-2011 年不同产品出口额的大小，点越大出口额越大；该图用 cytoscape3.2.1 绘制而成，layout 选择 edge-weighted spring embedded layout。

资料来源：作者根据中国海关数据库（2002-2011）计算

3. 公式（2）后面跟着的解释中有这样的表达，“由于企业 f 在前一年未出口过产品 i ，因此企业本身对该产品的出口规模为 0”。那么这个“由于”从何而来呢？此外公式（2）、（3）、（4）的解释还是需要完全的。例如 δ_{ij} 表示什么， $\sum_j(\text{export}_{c,i})$ 中 j 的含义呢？我想还是要作者给出明确的表达。目前的修改稿是分别加了一点点解释，但并未对公式中的各部分作出很明确的表达的。还有，公式（3）中 $\text{export}_{c,j}$ 是 c 城市 j 产品出口， $\text{export}_{f,j}$ 是 c 城市中的 f 企业的 j 产品出口，那么公式（3）的整体表述呢？整体而言，这里对公式（2）、（3）、（4）的解释是不够的。

回复：这种低级的失误应自我检讨。文中已经认真修改。具体修改如下，公式编号略有变化。

在获得“产品—产品”间的技术相关度后，本文通过识别企业所在的城市，计算出口集聚和企业相关生产能力。其中，出口集聚的测度中，根据产品之间认知距离的大小，又分为同产品集聚（ Agg ）、相关产品集聚（ Agg_R ）和不相关产品集聚（ Agg_UR ）。同产品集聚程度（ Agg ）计算公式如下：

$$Agg_{c,g,i} = \sum_{f \in c, f \neq g} (\text{export}_{f,i}) \quad (4)$$

其中， c 是企业 f 和 g 所在的城市， i 为产品。 $\text{export}_{f,i}$ 为城市 c 中企业 f 出口产品 i 的规模。同产品集聚效应是指企业 g 所在城市 c 的其他企业出口产品 i 的集聚程度 $Agg_{c,g,i}$

对企业出口扩展的影响。通过引入产品技术关联，将大于某一技术关联值的产品定义为具有技术关联的产品对，相关产品集聚程度（ Agg_R ）计算公式如下：

$$Agg_R_{c,g,i} = \sum_j (\text{export}_{c,j} \times \delta_{ij}) - \sum_j (\text{export}_{g,j} \times \delta_{ij}), \text{ 其中, } \delta_{ij} = \begin{cases} 1, & \phi_{ij} \geq 0.25 \\ 0, & \phi_{ij} < 0.25 \end{cases} \quad (5)$$

其中, $export_{c,j}$ 和 $export_{g,j}$ 分别为企业 g 所在城市 c 和企业 g 出口产品 j 的总规模。 ϕ_{ij} 为基于公式 (1) 计算的产品 i 和 j 之间的技术相关度。为区分相关产品和不相关产品集聚, 本文选择以 0.25 作为产品间技术相关度的临界值⁵。 ϕ_{ij} 大于等于 0.25, 则认为两产品在技术上相关, δ_{ij} 取 1, 否则 δ_{ij} 取 0。以 δ_{ij} 加权分别得到城市 c 中与产品 i 技术相关产品的出口总规模 ($\sum_j(export_{c,j} \times \delta_{ij})$) 和城市 c 中企业 g 与产品 i 技术相关产品的出口规模 ($\sum_j(export_{g,j} \times \delta_{ij})$), 两者相减得到企业 g 所在城市其他企业与产品 i 技术相关产品的出口集聚程度 $Agg_R_{c,g,i}$, 其对企业 g 出口产品扩展的影响即为相关产品集聚效应。同理, 识别技术关联度小于 0.25 的产品对, 以此计算不相关产品集聚程度 (Agg_UR), 计算公式如下:

$$Agg_UR_{c,g,i} = \sum_j[export_{c,j} \times (1 - \delta_{ij})] - \sum_j[export_{g,j} \times (1 - \delta_{ij})],$$

$$\text{其中, } \delta_{ij} = \begin{cases} 1, \phi_{ij} \geq 0.25 \\ 0, \phi_{ij} < 0.25 \end{cases} \quad (6)$$

所有符号的含义同上。不相关产品集聚 ($Agg_UR_{c,g,i}$) 对企业 g 出口扩展的影响即为不相关产品集聚效应。

4. 整体感觉, 仅仅单独列出公式 (6) 所表示的模型是不够的。毕竟在实证分析中经验模型选择有很多种形式, 或者是至少要把带交叉项的模型列出来, 并对两类模型分别用来检验溢出效应和企业相关生产能力做一个整体上的说明。这样可能会帮助阅读者更好地把握作者的整体思维。

回复: 文中增加了式 (7), 并对两类模型分别用来检验溢出效应和企业相关生产能力做一个整体上的说明, 使文章的整体性有了明显提高。

为验证假说 1, 本文将出口集聚分解为相同产品集聚和不同产品集聚, 后者基于技术相关度进一步分解为技术相关产品集聚和不相关产品集聚。为了考察三种类型的出口集聚对中国企业出口扩展边际的影响, 本文建立计量模型如 (1) 式所示。

$$Entry_{c,g,i,t} = \alpha_0 + \beta_1 Agg_{c,g,i,t-1} + \beta_2 Agg_R_{c,g,i,t-1} + \beta_3 Agg_UR_{c,g,i,t-1} + Control_{c,g,i,t-1} + \gamma_{g,i} + \gamma_{c,i} + \varepsilon_{c,g,i,t} \quad (1)$$

其中, 下标 c 、 g 、 i 、 t 分别表示区域、企业、产品和年份。 $Entry$ 是二元变量, 表示企业是否扩展产品 i , 扩展为 1, 否则为 0。 β_1 、 β_2 、 β_3 分别考察同产品集聚效应、相关产品集聚效应和不相关产品集聚效应, 所有解释变量均做取对数处理。 $Control$ 为企业—产品层面或企业层面的控制变量, 包括企业相关生产能力 ($CAPAB$)、企业出口总额 ($EXPORT$)、企业出口产品多样化水平 (DIV)、企业劳动生产率 ($PROV$) 和企业规模 ($SIZE$)。其中, 企业出口产品多样化水平为企业出口产品的种类数, 企业劳动生产率为单位劳动力的企业生产总值, 企业规模为企业从业人数, 企业出口总额和企业规模做取对数处理。

在 (1) 式的基础上, 本文引入企业相关生产能力 $CAPAB$ 与三个出口集聚变量的交互项,

⁵在产品相关度矩阵中, 有 1084 种产品, 587528 对关联数。当门槛值设定为 0.1 时, 所有产品都在产品相关网络中, 但关联数剩五分之一, 即 117685; 当门槛值设定为 0.2 时, 产品剩 822 个, 关联数剩 32330; 当门槛值设定为 0.3 时, 产品剩 534, 关联数剩 8077; 当门槛值设定为 0.5 时, 产品数仅剩 173。门槛值的选择标准既要尽量包括更多的产品和关联, 又要去掉那些不能代表有技术关联的弱联系。综上, 借鉴 Neffke et al. (2012) 以 0.25 为门槛的设置, 本文也将门槛值设为 0.25, 此时产品空间中有 660 个产品、16519 对关联对。后文也将使用 0.2 和 0.3 作为门槛值来进行稳健性检验。

构建计量模型来验证假说 2，即企业相关生产能力是否能提高企业在出口扩展过程中从出口集聚中的获益程度，如（2）式所示。

$$Entry_{c,g,i,t} = \alpha_0 + \beta_1 Agg_{c,g,i,t-1} + \beta_2 Agg_{-c,g,i,t-1} + \beta_3 Agg_{-c,g,i,t-1} UR_{c,g,i,t-1} + \beta_4 Agg_{c,g,i,t-1} \times CAPAB_{g,i,t-1} + \beta_5 Agg_{-c,g,i,t-1} \times CAPAB_{g,i,t-1} + \beta_6 Agg_{-c,g,i,t-1} \times CAPAB_{g,i,t-1} + Control + \gamma_{g,i} + \gamma_{c,i} + \varepsilon_{c,g,i,t} \quad (2)$$

其中，所有符号含义与（1）式一致。交叉项系数 β_4 、 β_5 、 β_6 分别考察企业能力对同产品集聚效应、区域相关产品集聚效应和区域不相关产品集聚效应的影响。为避免交叉项间的多重共线性问题，三个交叉项将被分别引入（2）式。

企业出口产品扩展的定义通常分为两年定义法或多年定义法。由于企业出口产品变动频繁，为了规避企业过去出口的影响以及识别企业对该产品的首次出口，本文采用多年定义法。企业 g 在 $t-1$ 年及其之前所有年份没有出口产品 i ，而在 t 年开始出口，则 $Entry_{g,i,t}$ 取值为 1，产品 i 则为企业 g 的新扩展产品；企业 g 在 t 年之前（包括 t 年）的所有年份未出口 i 产品，则 $Entry_{g,i,t}$ 取值为 0，产品 i 则为企业 g 的拟扩展产品。

5. “结论与讨论”的第一个段落，放在这里的话有些分散了研究主题。到了写研究结论阶段，还以提出问题的方式来表述，可能不是很妥当。建议直叙或者调整为在讨论之前的一段过渡性段落，但写作的思维要稍微转变，要跟讨论结合起来。

回复：谢谢您的建议，作者采用直叙对“结论与讨论”部分重新进行了修改，请见修改稿第六部分。

国际贸易领域在多产品企业模型上所进行的拓展，为企业内产品扩展边际的研究奠定了理论基础。已有研究虽然证实了地理距离刻画的出口集聚对企业出口动态和出口行为产生正向作用，却忽视了在通讯、传播技术迅猛发展的全球化时代，产品间认知距离对出口集聚效应的重要作用，更鲜有对集聚效应发挥条件的企业条件进行讨论。本文基于中国企业出口扩展的典型事实，讨论出口集聚以及企业相关生产能力对企业出口扩展的影响，并进一步做了稳健性、机制和异质性讨论。研究发现：①认知距离刻画的出口集聚显著提高了企业出口产品扩展边际，且认知距离越近，集聚效应越大。不同类型的出口集聚对企业出口扩展的影响存在显著差异，同产品集聚和相关产品集聚均表现为正效应，不相关产品集聚表现为负效应，即竞争效应。②不是每个企业都能“搭便车”，也就是说，集聚效应并不是平均地作用于每个企业，其发挥作用的方向和程度因企业储备的相关生产能力而异。高能力企业比低能力企业能更好地利用出口集聚效应来进行出口扩展。③影响机制方面，出口集聚效应不仅可以通过专业劳动力共享和中间投入品共享等供给侧渠道影响企业出口扩展，而且也可以通过市场信息溢出和已有出口经验共享等需求侧渠道来降低企业产品扩展的难度和风险。

在中国经济转型压力骤增和国际贸易战升级的双重背景下，提高中国企业出口产品多元化是中国调整出口结构、提升出口竞争力、促进经济持续增长的关键。产品多元化并不意味着企业可以向任意产品扩展，中国亟需在国际出口市场上打造自己的核心竞争力，而这从根本上需要依赖出口企业的不懈努力。近些年来，中国逐步脱下“加工车间”的帽子，部分出口企业开始掌握核心生产技术，在出口市场崭露头角，甚至在有些领域扮演着举足轻重的角色，然而与欧美国家的差距也给了我们继续努力的空间。为了鼓励中国出口企业不断创新，地方政府大力上马工业园区和自由贸易试验区等，效果却未达到预期。结合本文的发现，出口集聚，尤其是同产品集聚和相关产品集聚，始终是中国企业出口产品多元化的外源动力。针对地方政府：①地方政府应尽可能为企业打造技术关联紧密、上下游产业相互依托、有核心竞争力的空间集聚形态，充分发挥企业集聚效应，避免技术不相关的企业“扎堆”带来竞

争效应和资源重组的负面影响。②为实现这种空间集聚形态的形成,地方政府需要破除区域间的行政壁垒,开展区域协作,促进要素在区域间的自由流动,为提高产业空间集聚的技术相关度打下良好的市场基础。③地方政府除了提供补贴和税收优惠等产业政策外,更应为出口企业营造国际化、市场化、法治化的营商环境,提高企业间信息与知识交换的动力,激活集聚效应。针对企业而言,企业自身储备的生产能力才是企业出口扩展的内在动力,有针对性产业政策以提高企业核心竞争力,是中国市场经济可持续发展的微观基础。此外,中国企业虽然在不断的融入全球生产网络,但往往表现出显著的企业差异。外资企业由于多样化的资源获取途径在出口发展中有更大的灵活性和自主性,但其本地性较弱。因此,增加本地企业资源获取的途径,不断提升本地企业在企业集聚中的核心作用,将有利于增强本地经济持续增长的能力。

当然,本文的研究还较为初步,后续还有诸多有待完善的地方:①在刻画相关产品集聚和不相关产品集聚过程中,产品相关与不相关的门槛值还可以进一步确定。而且该门槛值的确定可以成为今后政府筛选产业或者企业进入专业工业园以更大程度地获取集聚效应的关键。②全球价值链时代,国际分工从产业间发展到了产品内,任何一个企业、城市乃至国家都只是产业链的一环,一件最终品的生产依赖于全球分工体系。本文只聚焦于企业在最终产品上的扩展,而企业扩展的产品究竟有多大比例来源于本地,来源于本地的部分究竟多大程度上受到集聚效应的影响需要进一步研究。③企业生产能力通过哪些渠道影响集聚效应,是企业生产能力发挥作用的关键,也是今后可以进一步讨论的方向。本文虽然对出口集聚效应发挥作用的企业条件进行了探讨,但由于无法获得企业各环节生产和交易的数据,没有做更为细致的检验,有待后续研究继续探讨。

参考文献:

- (1) 邓向荣,曹红. 产业升级路径选择:遵循抑或偏离比较优势——基于产品空间结构的实证分析[J]. 中国工业经济, 2016, (2):52-67.
- (2) 韩峰,柯善咨. 追踪我国制造业集聚的空间来源:基于马歇尔外部性与新经济地理的综合视角[J]. 管理世界, 2012, (10): 55-70.
- (3) 钱学锋,王胜,陈勇兵. 中国的多产品出口企业及其产品范围:事实与解释[J]. 管理世界, 2013, (1): 9-27.
- (4) 孙天阳,许和连,王海成. 产品关联、市场邻近与企业出口扩展边际[J]. 中国工业经济, 2018, (5): 24-42.
- (5) 吴群锋,杨汝岱. 网络与贸易:一个扩展引力模型研究框架[J]. 经济研究, 2019, (2): 84-101.
- (6) 杨汝岱,李艳. 移民网络与企业出口边界动态演变[J]. 经济研究, 2016, (3): 163-175.
- (7) Aitken B., G. H. Hanson, and A. E. Harrison. Spillovers, Foreign Investment, and Export Behavior[J]. *Journal of International Economics*, 1997, 43(1): 103-132.
- (8) Álvarez R., H. Faruq, and R. A. López. Is Previous Export Experience Important for New Exports[J]. *The Journal of Development Studies*, 2013, 49(3): 426-441.
- (9) Bishop P., and P. Gripiaios. Spatial Externalities, Relatedness and Sector Employment Growth in Great Britain[J]. *Regional Studies*, 2010, 44(4): 443-454.
- (10) Boschma R., A. Minondo, and M. Navarro. The Emergence of New Industries at the Regional Level in Spain: A Proximity Approach Based on Product Relatedness[J]. *Economic Geography*, 2013, 89(1): 29-51.
- (11) Danneels E. The Dynamics of Product Innovation and Firm Competences[J]. *Strategic Management Journal*, 2002, 23(12): 1095-1121.

- (12) Fernandes A. P., and H. Tang. Learning to Export from Neighbors[J]. *Journal of International Economics*, 2014, 94(1): 67-84.
- (13) Fleming L. Recombinant Uncertainty in Technological Search[J]. *Management science*, 2001, 47(1): 117-132.
- (14) Frenken K., F. Van Oort, and T. Verburg. Related Variety, Unrelated Variety and Regional Economic Growth[J]. *Regional Studies*, 2007, 41(5): 685-697.
- (15) Greenstone M., R. Hornbeck, and E. Moretti. Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings[J]. *Journal of Political Economy*, 2010, 118(3): 536-598.
- (16) Henderson V. Externalities and Industrial Development[J]. *Journal of Urban Economics*, 1997, 42(3): 449-470.
- (17) Hidalgo C. A., B. Klinger, and A. L. Barabási. The Product Space Conditions the Development of Nations[J]. *Science*, 2007, 317(5837): 482-487.
- (18) Jaffe A. B., M. Trajtenberg, and R. Henderson. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations[J]. *The Quarterly Journal of Economics*, 1993, 108(3): 577-598.
- (19) Koenig P. Agglomeration and the Export Decisions of French Firms[J]. *Journal of Urban Economics*, 2009, 66(3): 186-195.
- (20) Lemelin A. Relatedness in the Patterns of Interindustry Diversification[J]. *The Review of Economics and Statistics*, 1982, 64(4): 646-657.
- (21) Lo Turco A., and D. Maggioni. On Firms' Product Space Evolution: The Role of Firm and Local Product Relatedness[J]. *Journal of Economic Geography*, 2016, 16(5): 975-1006.
- (22) MacDonald J. M. R&D and the Directions of Diversification[J]. *The Review of Economics and Statistics*, 1985, 67(4): 583-590.
- (23) Neffke F. M. H., M. Henning, and R. Boschma. The Impact of Aging and Technological Relatedness on Agglomeration Externalities: A Survival Analysis[J]. *Journal of Economic Geography*, 2012, 12(2): 485-517.
- (24) Nooteboom B. Learning by Interaction: Absorptive Capacity, Cognitive Distance and Governance[J]. *Journal of Management and Governance*, 2000, 4(1-2): 69-92.
- (25) Penrose E. T. *The Theory of the Growth of the Firm*[M]. MA: Cambridge, 1959.
- (26) Poncet S., and F. S. D. Waldemar. Product Relatedness and Firm Exports in China[J]. *World Bank Economic Review*, 2015, 29(3):579-605.
- (27) Teece D. J., R. Rumelt, and G. Dosi. Understanding Corporate Coherence: Theory and Evidence[J]. *Journal of Economic Behavior & Organization*, 1994, 23(1): 1-30.
- (28) Teece D. J. Towards an Economic Theory of the Multiproduct Firm[J]. *Journal of Economic Behavior & Organization*, 1982, 3(1): 39-63.
- (29) Wang C. C., and G. C. S. Lin. Dynamics of Innovation in a Globalizing China: Regional Environment, Inter-Firm Relations and Firm Attributes[J]. *Journal of Economic Geography*, 2013, 13(3): 397-418.
- (30) Zhou, Y., S. Zhu, and C. He. Learning from Yourself or Learning from Neighbours: Knowledge Spillovers, Institutional Context and Firm Upgrading[J]. *Regional Studies*, 2019, 53(10): 1397-1409.