

语言网络与企业出口

——基于扩展引力视角的研究

王永进 孟珊珊

(南开大学 经济学院,天津 300071)

摘要:国家间的语言差异不仅会直接影响双边贸易,而且还会通过网络效应间接影响国际贸易。本文利用2000~2006年中国微观企业数据和双边语言距离数据构建了语言网络变量,考察了语言网络对中国企业出口的影响。研究表明:首先,语言网络显著促进了企业出口,出口目标国与企业上期出口国家之间的平均语言相似度越高,企业出口至该国的可能性越大。其次,语言网络对企业出口的影响主要通过降低交流成本和促进文化认同两种渠道实现。再次,异质性检验结果表明,高生产率企业、大规模企业、民营企业以及差异化产品出口时受益于语言网络的程度更高。最后,在经过内生性问题处理和一系列稳健性检验后,本文的研究结果依然成立。

关键词:语言网络;企业出口;扩展引力

中图分类号:F752.62 **文献标识码:**A **文章编号:**1003-5230(2020)06-0132-11

一、引言

语言和文化差异对各国的经济发展和贸易往来产生了广泛影响。在历史的长河中,人们通过语言传递和保存人类的文明成果。经过千百年的历史沉淀,世界各国形成了各自的民族语言,这也代表国家之间存在较大的语言和文化差异,而这种差异仿佛一种无形壁垒阻碍着各国的交流和贸易^{[1][2]}。具体而言,一国企业在国际市场搜寻购买者(或供应商)时,语言差异引起的文化偏好、沟通难度等将大大提高企业的前期搜寻成本,并在企业与合作者开展贸易活动时进一步加大企业的贸易成本。基于此,语言差异引起了诸多经济学者的关注,在现有文献中,学者们一般通过加入共同语言变量来研究或控制语言因素的影响。但是,这种二元变量的衡量标准忽略了国家语言的属性,因此,难以完全捕捉其对国际贸易的准确影响。与既有文献不同,本文将连续型语言距离变量与扩展引力相结合,基于“网络”视角从企业层面出发考察语言对国际贸易的影响,并对可能的影响渠道进行分析和验证。

收稿日期:2020-06-05

基金项目:国家自然科学基金面上项目“竞争中性原则与中国对外贸易利益:理论机制、实证检验与量化分析”(72073073);天津市研究生科研创新项目“出口网络与企业进口行为”(2019YJSB060)

作者简介:王永进(1983—),男,山东章丘人,南开大学经济学院教授,博士生导师;

孟珊珊(1992—),女,山东高唐人,南开大学经济学院博士生。

证。在当前全球经济低迷、国际贸易摩擦加剧的背景下,厘清语言网络与国际贸易之间的关系将为中国出口企业及政府有关部门提供战略和政策制定方面的新思路,因此本文的研究具有十分重要的理论和现实意义。

理论上,语言网络影响国际贸易的途径主要有两种:一是交流成本渠道,沟通作为语言最本质的功能,在国际贸易中扮演着重要的角色;二是文化和信任渠道,语言的形成往往和移民、种族、信任等词相互关联。为验证上述观点,本文结合2000~2006年中国工业企业数据库、中国海关数据库以及国家双边语言数据,从以下几个方面细致分析了语言网络对中国企业出口的影响。首先,我们基于语言距离数据和企业出口数据构建了本文的核心变量“语言网络”,使用Probit模型验证了语言网络对企业出口的促进作用;其次,为避免可能存在的内生性问题对基准回归结果产生影响,本文采用工具变量法和基于异方差的识别技术进一步准确识别了语言网络对企业出口的影响;再次,我们考察了语言网络影响企业出口的路径机制;最后,我们从四个方面进行了稳健性检验,确保本文研究结论的可靠性;此外,考虑到不同出口企业和出口产品受益于语言网络的程度存在明显差异,本文从四个方面进行了异质性检验。本文的主要结论如下:第一,语言网络显著增加了中国企业出口的可能性,目标国与企业上期出口国家之间的平均语言距离越小,企业出口至该国的可能性越大;第二,语言网络主要通过两个途径影响中国企业出口,一是交流成本渠道,二是文化认同渠道;第三,企业异质性检验结果表明,高生产率企业、大规模企业、民营企业以及差异化产品出口时受益于语言网络的程度更大;第四,在加入不同工具变量、更换核心解释变量测算方法、剔除部分特殊样本以及采用不同回归模型等一系列内生性问题处理和稳健性检验后,本文的研究结论依然成立。

本文余下的结构安排如下:第二部分为文献综述;第三部分为数据来源与相关变量说明;第四部分为计量模型设定与基准回归结果分析;第五部分做了进一步讨论;最后为本文的研究结论。

二、文献综述

(一)语言与国际贸易

大量证据表明,语言差异是阻碍国际贸易的壁垒之一。自20世纪80年代开始,Carr(1985)打开了探索语言与国际贸易关系的大门,他认为,共同语言在国际贸易中可以发挥与货币统一相似的作用——通过降低交易成本促进国际贸易^[3]。此后,考察语言对贸易的影响成为贸易研究领域的新热点。研究初始,多数文献专注于研究共同官方语言对双边贸易的影响,他们在引力模型框架内引入新的虚拟变量,若两国共享同一种官方语言,该变量为1,否则为0^[4]。Egger和Lassmann(2012)的研究表明,两国之间使用共同官方语言将使双边贸易流量提升33%~44%^[5],这也证实了Melitz(2008)早期有关语言影响国际贸易的研究结论^[1]。随着研究的不断深入,学者们开始意识到,共同官方语言的二元指标无法完全体现语言因素在国际贸易中的重要性。鉴于此,有关共同母语和共同口语如何影响国际贸易的文献开始不断涌现,例如Melitz和Toubal(2014)、Egger和Toubal(2016)以及Egger和Toubal(2018)^{[6][7][2]}。

除共同语言外,学者们还构建不同的语言距离指标对语言的重要性进行了考察。Lohammn(2011)构建了新的语言障碍指数,发现语言距离严重阻碍了两国之间的贸易往来^[8]。Melitz和Toubal(2014)除构建共同官方语言、共同口语和共同母语之外,还构建了基于民族语语言树和ASJP(Automated Similarity Judgment Program)等不同标准的语言相似度指标LP(Linguistic Proximity)来考察两国语言差异对双边贸易的影响^[6]。Lian等(2019)构建了语言摩擦指数分析语言和文化对跨国并购活动的影响^[9]。Konara(2020)使用语言连通性研究了语言在一国全球化过程中所发挥的作用^[10]。

(二)扩展引力与国际贸易

传统引力模型考察了地理距离、共同边界等双边因素对两国贸易的影响,扩展引力模型则侧重于分析出口目标国与企业或国家之前出口过的国家之间的地理距离、共同边界等双边因素在贸易中的作用。学者们认为,企业在选择出口目的地时,会更倾向于选择与其之前出口过的国家相似的目标国

市场^[11],这是由于出口至相似国家的市场进入成本相对较低,这种路径依赖被称为扩展引力^[12]。Chaney(2014)基于出口网络视角的研究认为,连续出口企业倾向于进入与其上一期出口国家之间平均距离更近的目标国市场^[13]。Defever等(2015)使用中国的企业数据研究了企业出口的空间特征,结果表明,若目标国市场与企业之前出口的国家之间存在共同边界,那么企业出口至该目标国的概率将增加两个百分点^[14]。在此基础上,Morales等(2019)的研究结果进一步证实,扩展引力确实对企业出口市场的进入成本有着重要影响^[12]。

(三)网络与国际贸易

根据既有文献,由于地理距离、文化制度差异、语言差异等非正式壁垒的存在,企业进行国际贸易时会面临较为严重的信息摩擦问题,高额的信息搜寻成本令众多出口企业望而却步。而网络作为国际贸易信息传播的重要载体,对国际贸易和对外投资的影响不言而喻^{[15][16]}。近些年,学者们从不同方面考察了网络在国际贸易中所发挥的作用,具体而言可以分为两类,一是企业外部网络,有关企业外部网络的研究已经非常丰富,例如移民网络^{[17][18]}、互联网网络^[19]、社会商业网络^{[20][21]};二是企业自生网络,相比外部网络,目前只有少数文献关注了企业自生网络,例如 Chaney(2014)对企业自生出口网络在未来出口中的重要性进行了研究^[13];Bernard等(2019)证实了产品网络结构对企业绩效的促进作用^[22]。

(四)文献评述及本文主要贡献

与现有文献相比,本文的边际贡献可能表现在以下几点:

其一,本文丰富了语言与贸易领域的研究。一方面,本文从扩展引力视角出发构建了语言网络变量,并研究了语言网络对国际贸易的影响。虽然经济学家已证实语言在国际贸易中扮演着重要角色,但既有文献要么将共同语言的二元变量加入引力模型作为研究双边贸易影响因素的控制变量,要么专注于研究共同官方语言、共同口语、共同母语或某种语言在国际贸易中的作用。基于此,本文借鉴Melitz和Toubal(2014)以及Spolaore和Wacziarg(2016)对语言距离的定义和测算^{[6][23]},使用双边语言相似性数据对语言网络的重要性进行了研究。另一方面,目前国际和国内有关语言与国际贸易关系的研究大多从国家层面展开,本文使用中国微观数据考察了语言网络对企业出口概率的影响。

其二,本文补充了国际贸易领域有关扩展引力因素的研究。Morales等(2019)检验了扩展引力在企业出口中发挥的作用,研究表明,当目标国与企业上期出口国家之间同时共享地理位置、共同语言以及相似收入水平时,企业进入目标国市场的7沉没成本最高减少90%之多^[12],由此也可见扩展引力的重要性。但其仍然使用二元变量衡量语言因素,势必会低估语言对国际贸易的影响。鉴于此,本文引入连续性语言距离变量,以目标国与企业上期出口国家之间的平均语言距离来测算语言网络,然后考察语言网络在国际贸易中的重要性。

其三,本文还扩展了网络与国际贸易关系的文献。随着学术界对信息摩擦如何影响国际贸易研究的不断深入,作为信息传播媒介的网络与国际贸易之间的关系也受到了学者们的广泛关注。但目前有关网络的研究主要集中于移民网络、互联网网络、社会与商业网络等企业外部网络,虽然 Chaney(2014)以及 Bernard等(2019)对企业自生网络进行了研究^{[13][22]},但均未涉及语言因素。鉴于此,本文将语言距离与网络相结合研究语言网络对国际贸易的影响,无疑丰富了网络与国际贸易领域的相关研究。

三、数据来源与核心变量构建

(一)数据来源

本文采用2000~2006年中国制造业企业数据作为经验研究样本,主要涉及企业层面数据和国家层面数据两部分。其中企业层面包括两套数据,一是中国工业企业数据库,该数据库涵盖了全部国有企业和规模以上非国有企业,详细记录了企业的各项个体信息,例如企业所在的地理位置、所有制属性、所属行业以及各项财务数据;二是中国海关进出口数据库,该数据库详细报告了企业进出口的

HS-8 位产品种类、价格、数量以及贸易方式等有效信息。国家层面数据主要来源于 CEPII 数据库，其中，双边引力变量来自 CEPII 的 gravity 数据库，语言距离数据来自 CEPII 的 language 数据库。另外，目标国的特征数据(人均 GDP)来自世界银行的 WDI 数据库。

由于构建语言网络指标的特殊需要，本文在将以上企业数据与国家数据进行匹配后，进一步对企业数据进行了连续性处理。此处的连续性定义为：在样本期间内至少连续出口两年及以上的企业。最终我们得到 17681 家企业出口至 161 个国家和地区的样本数据。

(二)核心指标构建

语言网络作为本文的核心解释变量，其数据的选取和构建的合理性将直接影响本文的估计结果。目前国际贸易领域的大多数文献使用两国之间是否拥有共同语言变量代表两国之间的语言差异，而简单的 0-1 变量可能无法准确表达两国之间的语言距离。因此，本文借鉴 Melitz 和 Touble(2014)以及 Spolaore 和 Wacziarg(2016)对语言相似性的两种测算方法：一种是基于语言树之间的差异，另一种是基于词汇统计^{[6][23]}。

第一种测算方法中语言树的概念源于分类学，学者将相似度较高的语言归为一类，例如，在民族语常用的分类中，法语和意大利语之间有四个共同的节点特征，这将作为衡量二者之间语言距离的基础^[6]。总而言之，两种语言之间的共同节点特征数量和语言距离是相对应的，例如，法语和意大利语两种语言与非印欧语言之间均无共同节点特征，故二者与非印欧语言之间的语言距离将远大于二者之间的语言距离。Spolaore 和 Wacziarg(2016)在 Fearon(2003)的数据基础上计算了两种语言的共同节点特征数量 CN(取值在[0, 15]范围内)，然后进行如下变换得到语言相似度变量(prox1)^{[23][24]}：

$$\text{prox1} = \sqrt{\frac{15 - \text{CN}}{15}} \quad (1)$$

第二种测算方法中的词汇统计为定量语言学的分支，这种方法根据常见单词是否同源(即源自同一祖先单词)对语言进行分类。具有多个同源单词的两种语言的距离往往比不具有同源词的语言更近^[23]，通过统计不同语言之间的同源词，并使用同源词所占百分比衡量语言之间的相似度。该方法计算得出的语言相似度(prox2)取值在[0, 0.92]范围内，样本均值约为 0.6。与第一种语言树测算方法相比，这种方法的重大优势在于其测算出的语言相似度为连续变量，而第一种方法测算出的是离散变量，显而易见，连续变量比离散变量更具有代表性。此外，本文在进行词汇统计和对比时使用了自动相似性判断方法(ASJP)，故也将第二种测算方法称为基于 ASJP 的测算方法。

在获得语言距离(语言相似度 prox)变量数据的基础上，本文参考 Chaney(2014)对网络的定义进一步构建语言网络指标(ling_net)^[13]，具体如下：

$$\text{ling_net} = \frac{\sum_{j'} 1 [\text{exporting}_{ij't-1} > 0] \text{prox}_{ij'}}{\sum_{j'} 1 [\text{exporting}_{ij't-1} > 0]} \quad (2)$$

式(2)中， $1 [\text{exporting}_{ij't-1} > 0]$ 为企业出口状态变量，若企业 i 在 $t-1$ 时期出口至 j' 国取值为 1，否则为 0； $\text{prox}_{ij'}$ 为企业出口目标国 j 与 j' 国之间的语言相似度。该变量的计算含义为企业潜在出口目标国 j 与企业上一期出口国家 j' 之间的平均语言相似度，该变量值越大，代表潜在目标国与企业已出口国家之间的语言相似度越高、距离越小，本文预测企业出口至该国家的概率越高。

鉴于语言距离的第二种测算方法比第一种更优，故本文在基准回归中将采用第二种语言相似度(prox2)，并在稳健性检验中替换为基于语言树的第一种语言相似度(prox1)来验证本文研究结论的一致性。

四、模型设定与基准估计结果

(一)计量模型设定

参考传统国际贸易文献的做法，本文仍然选择以引力模型为基础在企业层面进行概率模型的实证检验，模型具体如下：

$$\Pr(\text{exporting}_{ijt} > 0) = \Phi(\alpha \text{ling_net}_{ijt-1} + \beta X_{\text{CHN},j} + \gamma X_i + \eta X_j) \quad (3)$$

式(3)中, exporting_{ijt} 为企业 i 在 t 时期向 j 国的出口状态, 该变量为 0-1 变量, 企业出口时取值为 1, 不出口则为 0。 Φ 为标准正态分布的累计分布函数, 下标 i 为企业, j 为出口目标国, t 为时间, CHN 代表企业所在地中国。 $X_{\text{CHN},j}$ 表示中国与目标国之间的双边变量, 具体包括: 中国与目标国之间的地理距离 (Indist , 本文以使用人口加权后的两国首都之间的球面距离对数值来衡量)、是否具有共同边界 (border)、是否签订了自由贸易协定 (FTA) 以及双边语言相似度 (prox); X_i 表示企业层面的特征变量, 具体包括: 全要素生产率 ($\text{ln} \text{tfp}$, 本文采用 OP 方法进行测算, 取对数) 以及企业规模 ($\text{ln} \text{scale}$, 本文采用企业年末从业人数的对数值衡量); X_j 表示目标国国家层面特征变量, 即人均 GDP 的对数值。

(二) 基准回归结果

为检验语言网络对国际贸易的影响, 我们首先对式(3)进行了 Probit 回归, 基准回归结果如表 1 所示, 为便于解释回归系数的经济学含义, 本文汇报的 Probit 模型在基准回归以及后续其他检验中的估计系数均为平均边际效应。 其中, 第(1)列未添加任何控制变量, 仅加入核心解释变量 ling_net ; 第(2)列控制了引力变量, 具体包括国家层面的双边变量以及国家特征变量; 第(3)列进一步控制了企业层面的特征变量; 第(4)~(6)列分别在第(3)列的基础上进一步加入时间固定效应、行业固定效应以及时间和行业固定效应。 估计结果显示, 核心解释变量的估计系数与预期一致, 均为正且高度显著, 这表明语言网络对中国企业出口确实有一定的促进作用。 而且, 这一结果并没有因为固定效应的改变而发生显著变化, 同时加入各项控制变量、时间固定效应以及行业固定效应的回归结果(第(6)列)显示语言网络的回归系数为 0.3061, 该结果的经济学意义为: 语言网络变量增加 1%, 中国企业出口概率将增加 0.3061%。 企业语言网络的发展将为企业未来出口降低固定成本和沉没成本, 从而增加企业出口概率, 扩大企业出口范围。 此外, 其余控制变量的回归结果基本与预期保持一致, 囿于篇幅限制, 不再详述。

表 1 基准回归结果

	Probit			Probit_FE		
	(1)	(2)	(3)	(4)	(5)	(6)
ling_net	0.3737*** (0.0034)	0.3104*** (0.0030)	0.3054*** (0.0028)	0.3093*** (0.0029)	0.3013*** (0.0027)	0.3061*** (0.0028)
国家双边变量	NO	YES	YES	YES	YES	YES
目标国特征变量	NO	YES	YES	YES	YES	YES
企业特征变量	NO	NO	YES	YES	YES	YES
Sector FE	NO	NO	NO	NO	YES	YES
Year FE	NO	NO	NO	YES	NO	YES
Obs	4883571	4883571	4883571	4883571	4883571	4883571
Pseudo R ²	0.6666	0.1349	0.1453	0.1463	0.1532	0.1542

注: Probit 估计系数均为平均边际效应; ***, **, * 分别代表 1%、5%、10% 的统计显著水平; 括号内为在企业层面聚类的稳健标准误, 下表同。

(三) 内生性问题讨论与处理

内生性问题的来源之一是解释变量和被解释变量之间存在双向因果关系, 本文的被解释变量是企业层面的出口状态, 核心解释变量是滞后一期的语言网络。 我们认为本文不存在严重的逆向因果关系问题, 具体原因如下: 其一, 双边语言距离是由国家官方语言的各项特征经过不同方法的对比计算所得, 而一国语言往往是在该国千多年的历史沉淀中形成的, 短时间内受到贸易的影响可以说是微乎其微; 其二, 虽然本文构建了企业—国家—时间层面的语言网络变量, 但该变量本身与被解释变量是滞后一期的关系, 且其定义为出口目标国与企业之前出口国家之间的平均语言距离, 显而易见, 语言网络不太可能受到下一期企业出口状态的影响。

内生性问题的第二个来源是遗漏重要变量, 虽然我们认为本文不太可能存在逆向因果关系, 但是

仍然可能存在遗漏变量问题。为解决遗漏变量对本文估计结果可能产生的影响,我们在基准回归中逐步加入国家层面的双边变量、特征变量以及企业特征变量来控制由于这些因素的干扰产生的遗漏变量偏差。在此基础上,我们进一步加入时间固定效应和行业固定效应,试图控制时间和行业层面的因素。尽管我们加入不同的控制变量和固定效应解决遗漏变量问题,但不能保证我们已经完全排除了可能存在的影响语言网络和企业出口的遗漏变量。

为准确识别语言网络对企业出口的影响,我们试图通过两种方法来解决潜在的内生性问题:一是通过工具变量法,二是通过基于异方差的识别技术。具体如下所述:

1. 工具变量法

如上文所述,两国之间的语言距离由两国语言的共同特征和差异性决定,而一国的语言特征是在千百年的历史沉淀中积累形成的,因此同时影响世界各国语言而不影响贸易的历史原因并不容易寻找。在此情况下,本文首先参考高超等(2019)的方法,使用1950年世界各国的移民数据构造第一个工具变量^[25];然后借鉴经典文献的做法,以滞后两期的解释变量作为第二个工具变量^[26]。

高超等(2019)在研究方言与区域贸易时以清代和民国时期的省际双边移民作为方言的工具变量,他们认为历史移民会对当地语言产生决定性影响,同时历史上发生大规模移民的原因大多是灾荒、战乱等不可抗力因素,因此具有更好的外生性和随机性^[25]。在此基础上,本文选取了联合国移民数据库中1950年世界各国的历史移民数据作为语言的工具变量,囿于移民数据的可得性,最早的移民数据可追溯至1945年,但该年份的数据缺失严重,若使用该年数据进行回归会损失绝大多数样本,从而可能造成估计结果的偏差,因此本文选取了数据相对较丰富的1950年的移民数据来构造工具变量,鉴于1950年距今的时间差距已经很大,可以认为该阶段的移民会对各国的语言和文化产生深刻的影响,同时具有外生性即不对多年后中国企业的出口产生直接的影响。由于我们将1950年的双边移民作为双边语言距离的工具变量,所以我们按照构建语言网络的方式构建相似的移民变量来作为语言网络的工具变量:

$$imm_IV = \frac{\sum_{j'} 1 [\text{expoting}_{ij't-1} > 0] imm_{ij'}}{\sum_{j'} 1 [\text{exporting}_{ij't-1} > 0]}$$

其中 $imm_{ij'}$ 代表国家 j 与 j' 之间的移民数量。

表2第(1)~(4)列汇报了使用两种工具变量的估计结果,为保证工具变量和回归模型的合理性以及估计结果的一致性,此处分别进行IVprobit估计和两阶段最小二乘估计,同时本文在表2的倒数第二行汇报了工具变量回归中第一阶段的F值。其中,第(1)~(2)列为将1950年移民作为工具变量的结果,第(3)~(4)列为将语言网络滞后两期作为工具变量的结果。由表2可知,一方面,两种工具变量在不同回归方法下第一阶段的F值均远远高于Staiger和Stock(1997)在弱工具变量检验中提出的临界值^[27],表明本文选取的两个工具变量均不存在弱工具变量问题。另一方面,在加入工具变量后,语言网络的回归系数都显著为正,表明语言网络对企业出口有显著的促进作用,这与基准回归结果保持一致,而且在不同的工具变量以及不同的回归方法下该系数均未发生符号和显著性的变化,表明了本文基准回归结果的可信性。

表2 内生性问题处理回归结果

	工具变量法				基于异方差的识别技术
	1950年移民		语言网络滞后两期		基于异方差的识别技术
	IVprobit	2SLS	IVprobit	2SLS	
(1)	(2)	(3)	(4)	(5)	
ling_net	0.3713*** (0.0678)	3.1426*** (0.1898)	0.4923*** (0.0149)	0.9249*** (0.0240)	0.8618*** (0.0074)
Obs	641277	641277	1370432	1370432	4883571
F-value	2801.46	550.61	1875.85	192.06	
adj R ²		0.1291		0.5118	0.0726

注:回归中均控制了国家双边变量、目标国特征变量和企业特征变量以及行业、时间固定效应,如不做特殊说明,下表同。

2. 基于异方差的识别技术

除加入有效的工具变量外,本文还参考 Lewbel(2012)提出的基于异方差的识别技术来准确识别语言网络对企业出口的影响^[28]。基于异方差的识别技术可以在缺少有效工具变量时对内生性问题进行处理,也可以在有效工具变量的基础上做进一步的回归检验来提高工具变量估计的效率,而且该方法突破了工具变量法中有关排除性约束的条件限制,在使用过程中只需满足误差为异方差的条件即可。具体而言,基于异方差的识别方法可以概括为两阶段:第一阶段使用普通最小二乘法对内生被解释变量和外生变量 $Z(Z \in X$ 或者 $Z=X)$ 进行线性回归,并获得估计残差 $\hat{\epsilon}$;第二阶段使用由第一阶段得到的残差所构建的工具变量 $(Z-\bar{Z})\hat{\epsilon}$ 进行第二阶段回归,其中 \bar{Z} 为 Z 的均值。使用该方法的检验结果如表 2 第(5)列所示,结果显示,语言网络对企业出口的正向影响依然显著,这与基准回归结果以及使用两种工具变量的回归结果保持一致。

五、进一步讨论

本节进行以下三个方面的进一步讨论:其一,影响路径探讨,我们首先通过中介效应模型识别语言网络影响国际贸易的文化认同渠道,然后通过控制文化因素、缩小样本范围等方法进一步识别语言网络影响国际贸易的交流成本渠道;其二,稳健性检验,我们通过替换语言网络指标、替换被解释变量、剔除英语母语国家以及更换估计方法等不同方式对本文的研究结论进行稳健性检验;其三,异质性检验,我们根据企业异质性贸易理论,分析了语言网络对贸易的影响是否会随出口企业和出口产品的异质性特征而出现差异性。

(一)影响路径分析

根据现有文献的分析,语言影响国际贸易的渠道可以归纳为两种:其一,交流成本渠道,交流作为语言最本质的功能,相似的语言将提高沟通效率,降低交流成本,从而促进贸易^[29];其二,文化认同渠道,语言同时作为一种社会因素,常常被认为是文化的一部分,故语言网络也可通过文化渠道影响国际贸易^{[30][31]}。

1. 文化认同渠道

Egger 和 Lassmann(2015)认为语言是一种历史沉淀的产物^[30],语言相近可能会促成彼此的信任以及相似的文化,进而影响国际贸易。多年以来,文化对国际贸易的影响已被众多国内外研究所证实,那么语言是否通过文化渠道影响了企业层面的贸易呢?为回答此问题,本文将使用如下中介效应模型进行论证:

$$\Pr(\text{exporting}_{ijt} > 0) = \Phi(\alpha_0 \text{ling_net}_{ijt-1} + \alpha_1 X_{\text{CHN},j} + \alpha_2 X_i + \alpha_3 X_j) \quad (4)$$

$$\text{cul_net}_{ijt-1} = \beta_0 \text{ling_net}_{ijt-1} + \beta_1 X_{\text{CHN},j} + \beta_2 X_i + \beta_3 X_j + \epsilon_{ijt} \quad (5)$$

$$\Pr(\text{exporting}_{ijt} > 0) = \Phi(\gamma_0 \text{ling_net}_{ijt-1} + \gamma_1 \text{cul_net}_{ijt-1} + \gamma_2 X_{\text{CHN},j} + \gamma_3 X_i + \gamma_4 X_j) \quad (6)$$

式(5)中,cul_net_{ijt-1}为与语言网络对应构建的文化网络变量,该指标的构建同样参考 Chaney(2014)^[13],其中所使用的文化距离数据来自 Spolaore 和 Wacziarg(2016)^[23]。表 3 汇报了中介效应模型的回归结果,与式(4)~(6)相对应,第(1)~(3)列为相应模型的回归结果,为保证该结果的稳健性,本文使用线性概率模型再次进行了估计,结果见第(4)~(6)列。

表 3 第(1)列为语言网络影响企业出口的基准回归结果,语言网络变量的系数显著为正,表明语言网络提高了企业出口概率;第(2)列检验语言网络对文化网络的影响,语言网络的系数为负且高度显著,表明语言网络对文化网络的形成有一定的促进作用;第(3)列我们将核心解释变量和中介变量同时加入回归模型中,语言网络变量的系数显著为正,文化变量的系数显著为负,表明文化距离的增大将会降低企业出口的可能性。表 3 最后两行汇报了中介效应与中介效应占总效应之比,我们可以看到,中介效应的大小约为 0.0439,占基准回归系数的 17.10%。此外,在更换回归模型后,第(4)~(6)列的估计结果并未发生实质变化,证实了该中介效应的稳健性。

表 3

语言网络的文化认同渠道检验结果

估计方法	Probit	OLS	Probit	LPM	OLS	LMP
被解释变量	exporting	cul_net	exporting	exporting	cul_net	exporting
	(1)	(2)	(3)	(4)	(5)	(6)
ling_net	0.2570 *** (0.0023)	-2.5998 *** (0.0503)	0.8134 *** (0.0120)	1.0741 *** (0.0107)	-2.5998 *** (0.0503)	1.1744 *** (0.0143)
cul_net			-0.0169 *** (0.0003)			-0.0178 *** (0.0004)
Obs	4883571	4883571	4883571	4883571	4883571	4883571
Pseudo R ²	0.1510		0.1799			
adj R ²		0.0247		0.1074	0.0247	0.1143
中介效应		0.0439			0.0463	
中介效应/总效应		0.1710			0.0431	

2. 交流成本渠道

语言的本质是交流,语言距离的增大会导致企业出口固定成本和可变成本增加^[5],对国际贸易产生不利影响。现有文献发现,使用相同官方语言的国家之间贸易往来更加密切,语言距离会影响交流效率从而对两国之间的贸易流量产生影响。Fidrmuc 和 Fidrmuc(2016)进一步指出,在语言影响贸易的渠道中,交流渠道的作用要大于历史、文化和信任等渠道^[29]。

我们已经通过中介效应模型证实了语言网络影响企业出口的文化认同渠道,但由于交流成本难以准确衡量,我们无法继续使用中介效应模型对此进行检验,因此,本文将参照高超等(2019)的研究,选用排除法对交流成本渠道进行验证^[25]。一方面,我们在式(6)的文化渠道检验基础之上,进一步控制中国与目标国之间的文化距离 culdist,以此来控制文化渠道的影响;另一方面,为尽可能消除文化渠道的影响,我们继续选取与中国文化距离相近的目标国来缩小样本,样本数据中各国与中国的文化距离取值范围为[-23.74, 73.77],我们选取其中与中国文化距离在[-10, 10]范围内的国家进行了小样本回归。

两类方法的回归结果如表 4 所示:各列语言网络变量的系数与基准回归结果一致,依然显著为正,表明在控制文化渠道的干扰之后,语言网络仍然可以显著促进企业出口,这也证实了语言网络通过交流成本渠道对企业出口产生了影响。

表 4

语言网络的交流成本渠道检验结果

	控制文化		缩小样本	
	(1)	(2)	(3)	(4)
ling_net	0.3960 *** (0.0041)	0.3641 *** (0.0039)	0.4259 *** (0.0107)	0.3830 *** (0.0097)
cul_net	-0.0152 *** (0.0003)	-0.0144 *** (0.0003)	-0.0136 *** (0.0007)	-0.0166 *** (0.0007)
culdist	-0.0034 *** (0.0004)	-0.0026 *** (0.0006)	-0.2140 *** (0.0047)	-0.4795 *** (0.0429)
Obs	4883571	4883571	394632	394632
Peseudo R ²	0.0617	0.1127	0.1669	0.1909

注:此表第(1)(3)等奇数列未加入控制变量,第(2)(4)等偶数列加入国家双边变量、目标国特征变量以及企业特征变量等控制变量。下同。

(二)稳健性检验

为证实本文基准回归结果的稳健性,本文将进行以下稳健性检验:其一,替换语言网络指标。本文在基准回归中选用了第二种语言相似性测算方法,本节我们将使用第一种语言相似性来构建语言网络变量,利用该指标的回归结果显示,除语言网络的系数略微减小以外,其显著性和符号无其他显著变化,这表明在替换语言网络指标后,语言网络对企业出口的影响依然显著且稳健。其二,检验语言网络对出口额的影响。本文将模型被解释变量重新设定为企业出口额变量并再次进行回归。估计

结果表明,语言网络的系数显著为正,表明语言网络不仅影响企业决策,而且会增加企业的贸易流量,再次证实了语言网络对国际贸易的正向影响。其三,剔除母语为英语的目标国样本。英语作为目前的国际通用语言,以英语为母语的國家肯定会因此在国际贸易中受益,从而有可能使本文的估计结果产生偏误,因此我们进一步剔除了以英语为母语的目標国后进行了子样本回归,结果显示,语言网络的估计系数显著为正,再次证实了本文研究结论的稳健性。其四,使用线性概率模型进行检验。为排除估计方法对研究结果的影响,本文将 Probit 模型替换为线性概率模型进行了回归,得到的回归结果显示,本文的核心结论并不因估计方法的改变而发生变化。因篇幅所限,具体结果留存备索。

(三)异质性检验

1.生产率异质性

根据既有文献的研究结论,生产率较高的企业出口概率更大,原因是只有生产率足够高的企业才可以克服高额的贸易成本。为检验语言网络对贸易的影响是否会因生产率差异而发生变化,我们使用 OP 法测算企业全要素生产率,然后进一步将语言网络与企业全要素生产率的交叉项加入基准回归模型中,结果如表 5 第(1)~(2)列所示。首先,语言网络的系数与基准回归结果保持一致,再次证实了本文结论的稳健性;其次,交叉项的系数为正且高度显著,这表明语言网络在全要素生产率较高的企业出口过程中发挥了更大的促进作用。

表 5 异质性检验一

	生产率异质性		规模异质性	
	(1)	(2)	(3)	(4)
ling_net	0.3825 *** (0.0037)	0.3155 *** (0.0032)	0.0280 * (0.0137)	0.0338 * (0.0132)
ling_net * Intfp	0.0322 *** (0.0059)	0.0380 *** (0.0052)		
ling_net * Inscale			0.0656 *** (0.0028)	0.0655 *** (0.0027)
Obs	4883571	4883571	4883571	4883571
Peseudo R ²	0.0751	0.1538	0.0788	0.1480

2. 规模异质性

在企业异质性理论中,除企业生产率外,企业规模也备受关注。规模较大的企业可以受益于规模经济带来的成本降低,在国际贸易中更容易克服出口成本从而获取利润。本文推测,语言网络对规模较大企业的出口具有更大的促进作用。为验证该推测,我们继续将企业规模与语言网络的交叉项加入回归模型进行交互效应检验。表 5 第(3)~(4)列汇报了该检验的结果,语言网络和交叉项的估计系数均为正且在 1%水平上显著,这证实了我们之前的推测,也再次表明本文的研究结论具有良好的稳健性。

3. 所有制异质性

考虑到语言网络对企业出口的影响可能因企业所有制不同而有所差异,本文将样本企业分为国有企业、民营企业 and 外资企业进行分组回归。表 6 所示的估计结果显示,语言网络对三类企业的出口决策均产生了显著的正向影响,而且语言网络对企业出口的影响在民营企业中更胜一筹,外资企业其次,对国有企业影响最小。究其原因,在进行国际贸易时国有企业有其独特的对外交流优势,受语言影响较小;外资企业的母国本身就位于国外,这也形成了其天然的语言资源优势;而民营企业在贸易中受交流成本等的限制较大,因此语言网络对其的促进作用更为明显。

4. 产品差异化异质性

不同文化环境下消费者对产品的偏好可能不同,而且可变成本和固定成本的重要性可能也会随产品种类发生变化^[30]。基于此,本文根据 Rauch(1999)的分类将产品按照差异化程度分为三种:同质化产品、有参考价格的产品以及差异化产品^[32],相应的回归结果见表 7。我们可以发现,随着产品差异化程度的提高,语言网络的估计系数逐渐增加,这表明语言网络对企业出口的促进作用随产品差异化水平增加而相应增大。

表 6

异质性检验二

	国有企业		民营企业		外资企业	
	(1)	(2)	(3)	(4)	(5)	(6)
ling_net	0.3531 *** (0.0049)	0.3006 *** (0.0041)	0.4114 *** (0.0054)	0.3189 *** (0.0043)	0.3642 *** (0.0055)	0.3068 *** (0.0046)
Obs	2103156	2103156	2301354	2301354	1675476	1675476
Peseudo R ²	0.0749	0.1385	0.0774	0.1757	0.0747	0.1414

表 7

异质性检验三

	同质化产品		有参考价格的产品		差异化产品	
	(1)	(2)	(3)	(4)	(5)	(6)
ling_net	0.3734 *** (0.0218)	0.3218 *** (0.0181)	0.4122 *** (0.0042)	0.3302 *** (0.0034)	0.4422 *** (0.0067)	0.3660 *** (0.0055)
Obs	118000	118000	1685475	1685475	3623895	3623895
Peseudo R ²	0.1089	0.1902	0.0714	0.1530	0.0698	0.1523

六、结论

本文结合 2000~2006 年中国工业企业数据库、中国海关数据库以及国家双边语言数据,将连续型语言距离变量与扩展引力相结合,细致分析了语言网络对中国企业出口的影响。研究发现:第一,语言网络显著增加了中国企业出口的可能性,目标国与企业上期出口国家之间的平均语言距离越小,企业出口至该国的可能性越大;第二,语言网络主要通过提高文化认同感和降低交流成本来促进企业出口;第三,在加入不同工具变量、更换核心解释变量测算方法、剔除部分特殊样本以及使用不同回归模型等一系列内生性处理和稳健性检验后,本文的研究结论依然成立;第四,异质性检验结果表明,高生产率企业、大规模企业、民营企业以及差异化产品出口时受益于语言网络的程度更大。

本文的研究结论充分肯定了语言网络和企业出口经验在企业未来出口过程中所发挥的积极作用,从而为企业的出口市场选择提供了新思路。企业出口语言网络有助于降低企业与目标国之间的语言交流成本、提高沟通效率,从而降低企业未来出口市场的进入成本,出口企业尤其是连续出口企业可以在维系现有网络关系的同时继续扩大其出口语言网络的覆盖范围,如此良性循环将有利于进一步提高企业的出口绩效。长期以来,人们意识到了文化对国际贸易的影响,事实上语言因素在一国企业出口过程中所发挥的作用也不可忽视。基于此,在国家层面,我国政府可以加大教育经费投入,在高校中开设更加多元的语言专业和课程培养更多语言技术型人才;在企业层面,企业应该结合自身出口经验,在充分发挥出口语言网络优势的基础上,重点引进、培养多元化人才和现有语言网络未覆盖地区的语言专业型人才,为未来的市场扩张打下坚实基础,助力中国企业更好地走出去。此外,对于日韩等与中国文化、语言均相近且在世界范围内占据重要贸易地位的国家应给予更多关注,此类国家将在企业出口语言网络的构建和扩张中发挥积极的促进作用,使得企业在出口过程中可以更大程度地从出口语言网络中受益。

参考文献:

- [1] Melitz, J. Language and Foreign Trade[J]. *European Economic Review*, 2008, 52(4):667—699.
- [2] Egger, P. H., Toubal, F. Native Language and Acquired Language as Determinants of Product-level Trade [J]. *The World Economy*, 2018, 41(7):1833—1846.
- [3] Carr, J. Le Bilinguisme au Canada: L'usage Consacre-t-il L'anglais Monopole Naturel[C]//Vaillancourt, F. *Economie et langue* (Quebec: Conseil de la langue française). Montréal: Université de Montréal, 1985:27—37.
- [4] Eaton, J., Kortum, S. Technology, Geography, and Trade[J]. *Econometrica*, 2002, 70(5): 1741—1779.
- [5] Egger, P. H., Lassmann, A. The Language Effect in International Trade[J]. *Economics Letters*, 2012, 116

- [6] Melitz, J., Toubal, F. Native Language, Spoken Language, Translation and Trade[J]. *Journal of International Economics*, 2014, 93(2): 351—363.
- [7] Egger, P. H., Toubal, F. Common Spoken Languages and International Trade[C]//*The Palgrave Handbook of Economics and Language*. Paris: Université Paris1 Panthéon-Sorbonne, 2016: 263—289.
- [8] Lohmann, J. Do Language Barriers Affect Trade? [J]. *Economics Letters*, 2011, 110(1):159—162.
- [9] Lian, Z., Sun, W., Xie, D. et al. Cultural Difference and China's Cross-border M&As: Language Matters[J]. *International Review of Economics and Finance*, Available online 20 November 2019.
- [10] Konara, P. The Role of Language Connectedness in Reducing Home Bias in Trade, Investment, Information, and People Flows[J]. *Research in International Business and Finance*, 2020, 52(110118):1—16.
- [11] Evenett, S. J., Venables, A. J. Export Growth in Developing Countries: Market Entry and Bilateral Trade Flows[Z]. Mimeo, 2002.
- [12] Morales, E., Sheu, G., Zahler, A. Extended Gravity[J]. *The Review of Economic Studies*, 2019, 86(6): 2668—2712.
- [13] Chaney, T. The Network Structural in International Trade[J]. *American Economic Review*, 2014, 104(11): 3600—3634.
- [14] Defever, F., Heid, B., Larch, M. Spatial Exporter[J]. *Journal of International Economics*, 2015, 95(1): 145—156.
- [15] 吕越, 尉亚宁. 贸易网络地位与对外直接投资——基于1970~2017年跨国面板数据的经验证据[J]. *中南财经政法大学学报*, 2020, (5):113—124.
- [16] Allen, T. Information Frictions in Trade[J]. *Econometrica*, 2014, 82(6): 2041—2083.
- [17] Rauch, J. E., Trindade, V. Ethnic Chinese Networks in International Trade[J]. *Review of Economics and Statistics*, 2002, 84(1):116—130.
- [18] 蒙英华, 蔡宏波, 黄建忠. 移民网络对中国企业出口绩效的影响研究[J]. *管理世界*, 2015, (10):54—64.
- [19] 施炳展. 互联网与国际贸易——基于双边双向网址链接数据的经验分析[J]. *经济研究*, 2016, (5): 172—187.
- [20] Rauch, J. E. Business and Social Networks in International Trade[J]. *Journal of Economic Literature*, 2001, 39(4): 1177—1203.
- [21] Garmendia, A., Llano, C., Minondo, A., et al. Networks and the Disappearance of the International Home Bias[J]. *Economics Letters*, 2012, 116(2): 178—182.
- [22] Bernard, A., Moxnes, A., Saito, Y. Production Networks, Geography and Firm Performance[J]. *Journal of Political Economy*, 2019, 127(2): 639—688.
- [23] Spolaore, E., Wacziarg, R. Ancestry, Language and Culture[C]//*The Palgrave Handbook of Economics and Language*. London: Palgrave Macmillan, 2016: 174—211.
- [24] Fearon, J. Ethnic and Cultural Diversity by Country[J]. *Journal of Economic Growth*, 2003, 8(2): 195—222.
- [25] 高超, 黄玖立, 李坤望. 方言、移民史与区域间贸易[J]. *管理世界*, 2019, (2):43—57.
- [26] Egger, P. H., Erhardt, K., Lassmann, A. Immigration and Firms' Integration in International Production Networks[J]. *European Economic Review*, 2019, 111(1):1—34.
- [27] Staiger, D., Stock, J. H. Instrumental Variables Regression with Weak Instruments[J]. *Econometrica*, 1997, 65(3): 557—586.
- [28] Lewbel, A. Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models[J]. *Journal of Business and Economic Statistics*, 2012, 30(1):67—80.
- [29] Fidrmuc, J., Fidrmuc, J. Foreign Languages and Trade: Evidence from a Natural Experiment[J]. *Empirical Economics*, 2016, 50(1): 31—49.
- [30] Egger, P., Lassmann, A. The Causal Impact of Common Native Language on International Trade: Evidence from a Spatial Regression Discontinuity Design[J]. *Economic Journal*, 2015, 125(584): 699—745.
- [31] Guiso, L., Sapienza, P., Zingales, L. Cultural Biases in Economic Exchange? [J]. *Quarterly Journal of Economics*, 2009, 124(3):1095—1131.
- [32] Rauch, J. E. Networks versus Markets in International Trade[J]. *Journal of International Economics*, 1999, 48(1): 7—35.